STRUCTURE AND ACTIVITIES OF THE SACCHAROMYCES CEREVISIAE NUCLEAR RNA EXOSOME

A Dissertation

Presented to the Faculty of Weill Cornell Graduate School

Of Medical Sciences

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

John Charles Zinder

March 2018

STRUCTURE AND ACTIVITIES OF THE SACCHAROMYCES CEREVISIAE NUCLEAR RNA EXOSOME

John Charles Zinder, Ph.D

Cornell University, 2018

The eukaryotic RNA exosome is an essential and conserved protein complex that can degrade or process RNA substrates in the 3' to 5' direction. The nuclear RNA exosome includes a non-catalytic donut-shaped core (Exo9) that binds Dis3 (aka Rrp44) and the Rrp6/Rrp47 heterodimer to modulate their processive and distributive exoribonuclease activities, respectively. Additionally, cofactor proteins such as Mpp6 and the TRAMP and Ski complexes assist the exosome in RNA decay in different subcellular compartments.

Decades of study have revealed that the exosome acts on all classes of RNA in diverse model organisms and that the catalytic subunits primarily engage these substrates by first threading them through a prominent central channel in Exo9. Structural and biochemical studies have demonstrated that this channel is wide enough to permit single-stranded but not double-stranded RNA to enter, presenting an obstacle for degradation of structured substrates. This can be overcome by extension of the RNA 3' end, which can be accomplished by the polyadenylation and RNA helicase activities of the nuclear TRAMP complex. Recruitment of TRAMP to the exosome and the interplay between the various activities contained within these complexes remain unclear due to lack of rigorous biochemical characterization.

Contained within this dissertation are our efforts to investigate this important process, starting with structural characterization of the cofactor-less nuclear exosome (Exo9 plus Rrp6 and Dis3) in Chapter 1. Use of an engineered substrate enabled crystallization and X-ray structure determination of this complex, which uncovered

features in the non-catalytic core that modulate Dis3's activity in vitro. We also show that a 3' phosphate containing RNA cannot be trimmed by Rrp6 and instead is fully degraded by Dis3. Chapters 2 and 3 concern recruitment of Mtr4, the TRAMP complex RNA helicase, to the exosome. In Chapter 2, we solve the crystal structure of the nuclear exosome bound to the cofactor Mpp6. Biochemical characterization of the Mpp6-exosome revealed that Mpp6, along with Rrp6/Rrp47, physically tethers Mtr4 to the complex and enables ATP-dependent degradation of structured substrates. Chapter 3 details preliminary work on structurally characterizing a substrate-loaded Mtr4-exosome complex using cryo-electron microscopy.

BIOGRAPHICAL SKETCH

John Zinder was born in Ithaca, NY, where he attended Ithaca High School and subsequently Cornell University. There, he majored in Chemistry and began his research career in the laboratory of Drs Barbara Baird and David Holowka. His work in the Baird-Holowka lab concerned interactions between the cytoplasmic calcium channel Orai1 and the ER localized calcium sensor STIM1 in mast cells that occur after antigen stimulation, and he received the Merck Index Award for his efforts. After graduating with honors, he went on to the University of Colorado at Boulder to work in Dr. Charles McHenry's lab studying the roles of the dual replicative polymerases in *Bacillus subtilis*. John developed a passion for protein purification and enzymology while at Boulder, and after leaving with a master's in biochemistry he took those interests to York Avenue to pursue his doctorate in the Tri-Institutional Ph.D program in chemical biology (TPCB). John joined the laboratory of Dr. Christopher Lima in 2014, obtained his first crystals of the exosome shortly thereafter, and was immediately hooked on structural biology.

In addition to the Merck Index Award, John was on the dean's list while at Cornell, was the recipient of an NIH biophysics training grant and two graduate teaching awards at Boulder, and has been a speaker at the New York Structural Biology Discussion Group and the Cold Spring Harbor mRNA Processing meeting. John's research has been published in Molecular Cell, Genes and Development, Nature Communications, eLife, and Nucleic Acids Research.

In addition to science, John has played bass for over 20 years in nearly 10 different bands, on two different continents, and on 10 different albums. He also enjoys distance running and cooking pasta in his Park Slope apartment.

DEDICATION

This work is dedicated to the memory of Norton and Marilyn Zinder, who through perseverance, kindness, and unparalleled brilliance helped establish the incredible community of scientists on York Avenue that persists today.

ACKNOWLEDGEMNTS

One of the early lessons my father, a professor of microbiology, taught me is the value of community as it pertains to being a scientist, undoubtedly a lesson that was impressed upon him by his father, whose lab members were affectionately known as the ‘Zinder Kinder.’ I would like to acknowledge him for teaching me this valuable lesson among many, and take this opportunity to further acknowledge a small fraction of the rest of the community that helped in my maturation.

To my mother, for feeding me, driving me around, indulging my middle school rockstar fantasies, and for teaching me to be kind to others by default. And to my brothers for fighting with me but not killing me when we were young and for always being there to talk now that we’re not.

To my partner Julia Brown, who has seen the emotional rollercoaster that experimental science can put one through and served as a source of calm and reason throughout.

To coaches Jason Trumble, Rich Bernstein, and in memory of Dan Fravil: for pushing me further than I knew I could be pushed, and for seeing my potential in spite of my own doubts.

To my grade school teachers, especially Mr. Demo, Mrs. Ged, Mrs. Lee, Mr. Anderson, Mrs. Ridenour, Mr. Drix, Mrs. Teukolsky and Mr. McMahon: all of whom made going to class something to look forward to.

To Drs. Barbara Baird and David Holowka, for allowing me as an undergrad to goof around in their lab, even if it was mostly to get a reliable dog sitter.

To Dr. Robert Batey, for teaching me biochemistry from the ground up, seeding my interest in RNA structural biology, and vouching for me when I needed to find a new graduate program. And to Dr. Scott Blanchard, for carefully considering my

application to said new graduate program, which undoubtedly facilitated my getting an offer.

To Dr. Charles McHenry, whose patience knows no bounds, who taught me how to stand on my own two feet in the lab, and who impressed upon me the value of a clean prep.

To my friends and colleagues at Boulder, including Zoe Dancy, Dr. Ely Porter, Jeff Swan, Dr. Joan Marcano, Kendra Krueger, Dr. Jacob Polaski, Katie Spahr and Chalie Nevarez: for making my time out there so incredibly memorable.

To my classmates in TPCB and the other members of the Lima lab for always being willing to grab a beer and talk science. And to the Tri-I neighborhood for always having a happy hour with free beer in which to do so.

To Dr. Elizabeth Wasmuth, for being an excellent and attentive mentor as well as a loyal friend. I am reminded daily that I stand on the shoulder of giants, and I'm thankful that this particular one is of the gentle-giant variety.

And finally, to Dr. Chris Lima, for succeeding in the incredibly demanding task of mentoring me. I entered the lab thinking I knew everything, a dangerous mindset for any discipline, and his guidance has opened my eyes to just how much there is to know. I will always strive to be a savvier, smarter scientist because of my time in his lab and to simultaneously see the big-picture and the small-picture even a fraction as well as he does. As I move on to the next chapter, I'm confident that my training with Chris as my mentor has prepared me well, and look forward to whatever challenges I may face in the future.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

RNA decay in eukaryotes

RNA decay is an essential cellular process that regulates gene expression, generates mature RNAs from precursors, and monitors RNA quality in all organisms. RNAs can be degraded in either of three ways: from 5' end towards the 3' end by 5'-3' exoribonucleases, from 3' end towards the 5' end by 3'-5' exoribonucleases, or by cleavage in the middle by endoribonucleases. Every organism has evolved an ensemble of RNA degrading enzymes (RNases) possessing unique enzymatic characteristics (e.g. decay rate, decay processivity, nucleotide specificity etc.), protein-protein interaction networks, and subcellular localizations that are tailored to the particular needs of that organism. In spite of this tremendous diversity, several themes of RNA decay have emerged that seem to be generally applicable. Most notably, decay rates by many RNases are affected by the proteins bound to that particular RNA, and activities of many exoribonucleases are regulated based on specific chemical or structural features of the ends from which they are degrading (Belair et al., 2017; Kilchert et al., 2016). These types of restrictions ensure that RNA decay can be regulated, often enabling for targeting of aberrant transcripts over healthy ones (quality control) or unneeded regulatory genes over housekeeping ones (gene regulation). Furthermore, functional RNAs are rarely transcribed in their fully mature form, often requiring many steps of endo- and exonucleolytic degradation to ensure they achieve their proper mature form. While this generation of excess RNA may seem wasteful at first glance, it can provide an effective check of the fidelity of transcription, RNA folding, protein association, and post-transcriptional modifications of diverse RNAs. A prototypical example of this is tRNA splicing, where the pre-

tRNA contains an intron that is excised by the essential Sen2 and Sen34 endonucleases in yeast (Ho et al., 1990). These nucleases associate with the essential Sen54 protein, which binds the acceptor stem and functions as a 'ruler' to position the endonuclease active sites precisely at the boundaries of the intron. This ensures that only properly folded tRNAs will have their introns removed, and misfolded tRNAs will be retain them and be targeted for degradation (Trotta et al., 1997).

Pervasive transcription of eukaryotic genomes generates a great diversity of RNAs (ENCODE Project Consortium, 2012), some of which go on to carry out specific functions but many of which are targeted by adaptor proteins for rapid degradation by the RNA decay machinery (Belair et al., 2017; Kilchert et al., 2016). Examples of the latter category include cryptic unstable transcripts (CUTs), promoter upstream transcripts (PROMPTs), and enhancer RNAs (eRNAs) (Pefanis et al., 2015; Preker et al., 2008; Wyers et al., 2005). Whether these RNAs result from transcriptional noise or can be stabilized under certain conditions to carry out discrete functions is still a subject of ongoing research. Stable transcripts, by contrast, are thought to escape degradation via co-transcriptional recruitment of stabilizing factors (e.g. capping enzymes, the spliceosome). Some healthy, functional RNAs can also be degraded at faster rates than others, with housekeeping genes tending to have slower decay rates than regulatory genes (Hanson and Coller, 2018; Lima et al., 2017; Presnyak et al., 2015; Yang et al., 2003). This both limits the steady-state expression levels of regulatory proteins and enables cells to rapidly alter their transcriptional program in response to changes in the extracellular environment.

In this introduction, we will first give a brief survey of conserved exonucleases other than the exosome and their RNA targets, structures, and roles within the cell. Then we will move on to discuss the exosome and its cofactors, including structures, biological functions, and means of selective RNA targeting.

XRN1 and XRN2: the major eukaryotic 5'-3' exonucleases

In eukaryotes, 5'-3' exonucleolytic degradation of RNAs is primarily carried out by the structurally related Xrn1 and Xrn2 (aka Rat1) nucleases in the cytoplasm and nucleus, respectively. These enzymes are highly processive RNA degradation machines, but they cannot act on capped or recessed 5' ends, thus restricting their activities on properly processed and structured RNAs (Akiyama et al., 2016).

Xrn1 is the primary source of mRNA turnover in the cytoplasm of *S. cerevisiae* (Anderson and Parker, 1998), and it participates in the translation-dependent RNA decay pathways non-stop-, no-go-, and nonsense-mediated decay (NSD, NGD, NMD) (Łabno et al., 2016a). Additionally, *S. cerevisiae* Xrn1 targets a subset of unstable non-coding RNAs (ncRNAs) known as Xrn1-sensitive unstable transcripts (XUTs) for rapid degradation, thus preventing their accumulation (van Dijk et al., 2011). The structure of *Drosophila melonogaster* Xrn1 bound to a small fragment of 5' phosphate containing DNA (to prevent turnover during crystallization) revealed that the 5' terminus of the substrate resided in a conserved basic patch at the end of a narrow cleft which could accommodate approximately 4 nucleotides (nt) of single-stranded nucleic acid (Jinek et al., 2011). This situates the penultimate nt perfectly within the hydrolytic active site and explains why a 5' OH at the 5' end but not modifications significantly bulkier than a phosphate, such as a triphosphate or m7GpppG cap, can be degraded by the enzyme.

During transcription, Xrn2 recognizes the 5' end of downstream RNAs that have been cleaved after polyadenylation or co-transcriptional cleavage site and rapidly degrades them, which in turn stimulates dissociation of RNA polymerase from the template in a so-called 'torpedo' mechanism (Proudfoot, 1989; West et al., 2004). Additionally Xrn2 functions in 5' end processing of rRNA and snoRNAs in the nucleolus (Gasse et al., 2015; Nagarajan et al., 2013; Łabno et al., 2016a). The

structure of Rat1 bound to its partner Rai1 from *Schizosaccharomyces pombe* resembles the active site architecture of *Drosophila* Xrn1 but highlighted a pyrophosphohydrolase activity of Rai1 (Xiang et al., 2009). This activity can remove the beta and gamma phosphate from 5' triphosphate RNAs in vitro, leaving behind Xrn2's preferred substrate. While the interaction between Rai1 and Xrn2 is poorly conserved in other organisms, this suggests a strategy by which Xrn2 can degrade aberrantly uncapped or viral RNAs (Nagarajan et al., 2013; Xiang et al., 2009). In vertebrates, three proteins contain Xrn2-binding domains (XTBD) that bind and stabilize Xrn2 via the same conserved interface, suggesting the existence of multiple heterodimeric Xrn2 complexes in these organisms (Miki et al., 2014; Richter et al., 2016).

Other major eukaryotic 3' to 5' exonucleases

Eukaryotic genomes encode a variety of distributive 3'-5' exoribonucleases with specificity towards poly(A) RNA collectively known as deadenylases. In many cases mRNA deadenylation is thought to precede decapping and full degradation, underscoring the importance of these enzymes in RNA metabolism. The two most prominent among the deadenylases are the Pan2/3 and Ccr4-Not complexes (Collart and Panasenko, 2012; Wolf and Passmore, 2014; Łabno et al., 2016a). Pan2/3 is a predominantly cytoplasmic complex that consists of an asymmetric dimer of Pan3 proteins bound to a single molecule of Pan2 exonuclease (Jonas et al., 2014; Schäfer et al., 2014; Wolf et al., 2014). Because Pan2/3 can trim but not totally remove poly(A) tails and is stimulated by the presence of poly(A)-binding protein (PABP), it is thought that this complex initiates poly(A) tail removal (Lowell et al., 1992; Uchida et al., 2004; Łabno et al., 2016a). Ccr4-Not is a large multiprotein complex found in both the nucleus and cytoplasm of cells that contains two separate exonuclease subunits:

Ccr4 and Caf1/Pop2 (Collart and Panasenko, 2012; Stowell et al., 2016). Because Ccr4-Not can completely remove poly(A) tails and is inhibited by poly(A) binding protein (PABP) (Stowell et al., 2016; Tucker et al., 2002), fewer of which would be bound to shorter tails, it is thought to finish the deadenylation that was initiated by Pan2/3. Recently, it was shown that the most abundant *Caenorhabditis elegans* mRNAs contained approximately 30 or 60 nt poly(A) tails, which roughly corresponds to only one or two PABPs bound (Lima et al., 2017). Because initial poly(A) tail length is thought to be the same for all mRNAs, the authors suggest a role of deadenylases in pruning of these RNAs to achieve this apparently optimal poly(A) tail length on abundant RNAs. They also note that these abundant, short-tailed RNAs are very efficiently translated, have highly optimized codon content, and relatively long half-lives, suggesting crosstalk between the translation, deadenylation, and RNA decay pathways (Lima et al., 2017; Presnyak et al., 2015).

In addition to polyadenylation by poly(A) polymerases, RNA 3' ends of certain transcripts can be uridylated by terminal uridyl transfereases (TUTases) and targeted for degradation without the need for deadenylation. Dis3l2, a paralog of human Dis3 that does not associate with the exosome, is a processive 3'-5' exoribonuclease that targets 3' poly(U) containing RNAs for full degradation (Malecki et al., 2013; Viegas et al., 2015). Notably, both Dis3l2 and TUTases are present in diverse eukaryotes but absent in *S. cerevisiae* (Viegas et al., 2015; Łabno et al., 2016a). In higher eukaryotes, TUTases also add poly(U) tracts to mRNAs after deadenylation, after which point they are acted upon by both 5'-3' and 3'-5' decay pathways (Lim et al., 2014).

Genetic analysis of individuals with Perlman's overgrowth syndrome revealed that Dis3l2 was lacking exons critical for activity, and subsequent analysis of Dis3l2 knockdown cells in culture revealed a dysregulation of mitotic control (Astuti et al., 2012). Soon after this discovery, Dis3l2's activity was shown to play an important role

in the surveillance of *let-7* family micro RNAs. The pluripotency factor LIN28 recognizes *let-7* precursors (pre-*let7*) and recruits the TUTases Zcchc11 or Zcchc6 (aka TUT4 and TUT7, respectively), which add a ~14 nt poly(U) tract to allow degradation by Dis3l2 (Chang et al., 2013; Heo et al., 2009; Thornton et al., 2012; Ustianenko et al., 2013). Transcriptome-wide studies have connected Dis3l2's activity to surveillance of a wide variety of cytoplasmic ncRNAs in a uridylation-dependent manner (Pirouz et al., 2016; Łabno et al., 2016b). In *Drosophila*, the connection between 3' uridylation and Dis3l2-dependent decay is taken a step further: DmDis3l2 physically interacts with the TUTase Tailor to form the so-called TRUMP complex, which can degrade a wide variety of structured substrates in vitro and in vivo (Reimão Pinto et al., 2016). Structural and biochemical studies of mammalian Dis3l2 revealed base-specific contacts that underlie poly(U) recognition and an RNA path that explains the preference of the enzyme for RNAs containing 3' poly(U) tracts of ~14 nt or longer (Faehnle et al., 2014).

The eukaryotic RNA exosome

The eukaryotic RNA exosome is a conserved multi-subunit protein complex that catalyzes 3'-5' processing or degradation of a vast array of different RNA substrates (Januszyk and Lima, 2014; Kilchert et al., 2016). Since its discovery as a key factor involved in 3' processing of rRNAs during ribosome biogenesis in budding yeast (Mitchell et al., 1997), transcriptome-wide analyses in diverse eukaryotic model systems revealed that the RNA exosome contributes to the processing and/or degradation of every known class of RNA (Chekanova et al., 2007; Pefanis et al., 2014; Schneider et al., 2012).

Nuclear and cytoplasmic forms of the RNA exosome are defined by unique subunit compositions that interact with distinct cofactors in these subcellular

compartments (Table 1 and Figure 1). In the cytoplasm of *S. cerevisiae*, the exosome includes a nine-subunit core (Exo9) that interacts with Dis3 (aka Rrp44) to form a ten-subunit complex ($Exo10^{Dis3}$). The Exo9 core lacks catalytic activity while Dis3 catalyzes endoribonuclease (endo) and processive 3'-5' exoribonuclease (exo) activities (Dziembowski et al., 2007; Lebreton et al., 2008; Liu et al., 2006). Although redundant with cytoplasmic 5'-3' decay pathways (Anderson and Parker, 1998), $Exo10^{Dis3}$ contributes to translation-dependent mRNA surveillance pathways NSD, NMD, and NGD (Łabno et al., 2016a). All ten genes encoding subunits of $Exo10^{Dis3}$ are essential for viability in yeast (Brouwer et al., 2001; Mitchell et al., 1997). While *dis3* alleles that disrupt its endo activity bear few phenotypic defects, mutations that disrupt its exo activity result in slow-growth, and mutations that disrupt both activities result in a synthetic growth defects or inviability (Lebreton et al., 2008; Schaeffer et al., 2009; Schneider et al., 2008). In the nucleus, $Exo10^{Dis3}$ associates with a distributive 3' to 5' exoribonuclease Rrp6 (aka EXOSC10 and PM/Scl-100) and its obligate binding partner C1D (aka Rrp47) to form a 12-component complex ($Exo12^{Dis3/Rrp6/C1D}$) (Briggs et al., 1998; Feigenbutz et al., 2013a). While Rrp6 is not essential, Δ*rrp6* strains display a slow-growth phenotype, temperature sensitivity, and RNA processing defects (Allmang et al., 1999a; 1999b; Briggs et al., 1998).

Subunit compositions of nuclear and cytoplasmic RNA exosomes from human resemble yeast, with some notable differences. For instance, humans encode two exosome associated Dis3 enzymes that associate with the core, Dis3 and Dis3l, that localize to the nucleus and cytoplasm, respectively (Staals et al., 2010; Tomecki et al., 2010) (Figure 1). Similar to yeast, human Dis3 and Dis3l possesses exo activity, though only Dis3 has an intact endo catalytic site. Dis3 is excluded from the nucleolus in human cells while Rrp6, C1D, and Mpp6 are localized to the nucleus and enriched in the nucleolus (Schilders et al., 2007; Staals et al., 2010; Tomecki et al., 2010),

Table 1: RNA exosome and cofactors

	Human Gene	Protein Name(s)	Localization
S1/KH cap	*EXOSC1*	Csl4, Ski4	Nucleus/cytoplasm
	EXOSC2	Rrp4	Nucleus/cytoplasm
	EXOSC3	Rrp40	Nucleus/cytoplasm
PH-like ring	*EXOSC4*	Rrp41, Ski6	Nucleus/cytoplasm
	EXOSC5	Rrp46	Nucleus/cytoplasm
	EXOSC6	Mtr3	Nucleus/cytoplasm
	EXOSC7	Rrp42	Nucleus/cytoplasm
	EXOSC8	Rrp43	Nucleus/cytoplasm
	EXOSC9	Rrp45, PM/Scl-75	Nucleus/cytoplasm
Distributive 3'-5' exoribonuclease	*EXOSC10*	Rrp6, PM/Scl-100	Nucleus
Processive 3'-5' exoribonuclease, endonuclease	*DIS3*	Dis3, Rrp44	Nucleus (yeast), Nucleoplasm (human)
Processive 3'-5' exoribonuclease	*DIS3L1*	Dis3l1	Cytoplasm (human)
Exosome cofactor	*C1D*	C1D, Rrp47, Lrp1	Nucleus
Exosome cofactor	*MPP6*	Mpp6	Nucleus
Exosome cofactor	*HBS1L*	HBS1L (isoform 3, human), Ski7 (yeast)	Cytoplasm
Ski complex	*SKIV2L*	Ski2 helicase	Cytoplasm
	TTC37	Ski3	Cytoplasm
	WDR61	Ski8	Cytoplasm
TRAMP complex	*MTR4*	Mtr4 helicase	Nucleus
	ZCCHC7	Air1/2 (yeast), Air1 (human)	Nucleus (yeast), Nucleolus (human)
	PAPD5	Trf4/5 (yeast), PAPD5/Trf4-2 (human)	Nucleus (yeast), Nucleolus (human)

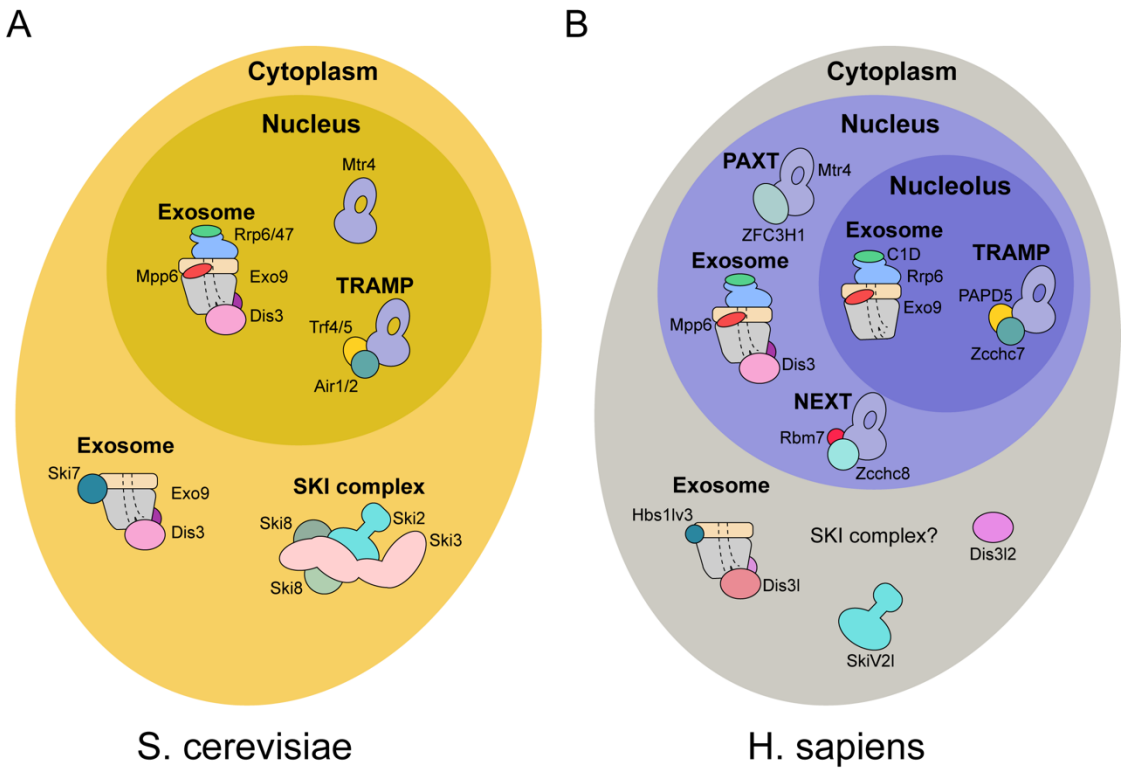


Figure 1. Localization and subunit composition of the exosome and related complexes. Schematic representations of the exosome composition and major exosome-related cofactors in the different compartments of *S. cerevisiae* (A) and human (B) cells. Cartoons were designed to reflect available structural, biochemical, and genetic data.

suggesting that the nuclear RNA exosome in humans includes Exo9, Dis3, Mpp6, and Rrp6/C1D (Exo13$^{Dis3/Mpp6/Rrp6/C1D}$) and that a nucleolar exosome may include Exo9, Mpp6, and Rrp6 as the only nuclease, presumably associated with C1D (Exo12$^{Mpp6/Rrp6/C1D}$). Mammalian cells lacking Dis3 cannot grow, and mutations that disrupt both its exo and endo activities are synthetic lethal in HeLa cells, indicating that Dis3 activities are not fully redundant with other RNA decay pathways (Tomecki et al., 2014).

RNA exosome structure, activities, and RNA paths to enzymatic subunits

While the structure of the human Exo9 core was obtained more than a decade ago, more recent crystal and cryo-electron microscopy (cryo-EM) structures have revealed architectures for intact yeast cytoplasmic and nuclear RNA exosomes in complex with RNA substrates and co-factors. Combined with biochemical and genetic studies, these structures illuminate roles for the non-catalytic core in modulating the activities of the associated ribonucleases and the impact of RNA path selection with respect to the fate of RNA substrates.

RNA exosome core and catalytic subunits

The Exo9 core includes a ring of six proteins (Rrp41, Rrp42, Rrp43, Rrp45, Rrp46 and Mtr3; collectively known as the PH-like ring) that are structurally homologous to the bacterial phosphorolytic exoribonuclease RNase PH, which consumes phosphate to degrade RNAs and releases NDPs as products. This ring is capped by a ring of three proteins with S1 and K-homology (KH) domains (Rrp4, Rrp40, and Csl4, which lacks a KH domain; the S1/KH cap). Together, Exo9 forms a non-catalytic donut shaped complex with a prominent central channel that is wide

enough to accommodate single-stranded RNA (ssRNA) (Liu et el., 2006; Dziembowski et al., 2007) (Figure 2 and Table 1). While the architecture of Exo9 closely resembles those of bacterial PNPase and the archaeal exosome, an important difference is that the bacterial and archaeal complexes contain phosphorolytic activity and Exo9 is devoid of catalytic activity (Dziembowski et al., 2007; Januszyk and Lima, 2014; Liu et al., 2006) (Figure 2A). A notable exception is the exosome core of *Arabidopsis thaliana*, which has recently been shown to contain phosphorolytic exonuclease activity in the Rrp41 subunit of its Exo9 (Sikorska et al., 2017). Inactivation of this activity resulted in specific rRNA processing defects, suggesting that it collaborates with the hydrolytic exonucleases within the nuclear complex. Phylogenetic analysis of the Rrp41 protein supported that this activity operates throughout the green lineage and suggests an ancestral eukaryotic RNA exosome that retained its phosphorolytic activity.

Catalytic subunits of the RNA exosome include Rrp6 and at least one isoform of Dis3. Dis3 and Dis3l (in the cytoplasm of higher eukaryotes) include an active site that catalyzes processive Mg^{2+}-dependent hydrolytic (meaning the enzymes consume water to degrade RNAs, releasing NMPs as products) 3' to 5' exoribonuclease activity (Dziembowski et al., 2007; Makino et al., 2013b; Staals et al., 2010; Tomecki et al., 2010). The exoribonuclease module of Dis3 encompasses a two cold-shock (CS) domain, a central catalytic RNB domain, and a C-terminal S1 domain (Figure 3A). Dis3's domain architecture and structure resemble those of bacterial RNAse II and RNAse R, as well as mammalian Dis3l2 (Chu et al., 2017; Faehnle et al., 2014; Frazão et al., 2006; Lorentzen et al., 2008) (Figure 3). The RNA path to the active site of ScDis3 is nearly identical to that of *Mus musculus* Dis3l2 and *E. coli* RNase II until the 6^{th} nt. After this point, it turns to enter a channel between CSD1 and the S1 domain in ScDis3 while it turns the opposite way to go between CSD2 and the S1 domain of

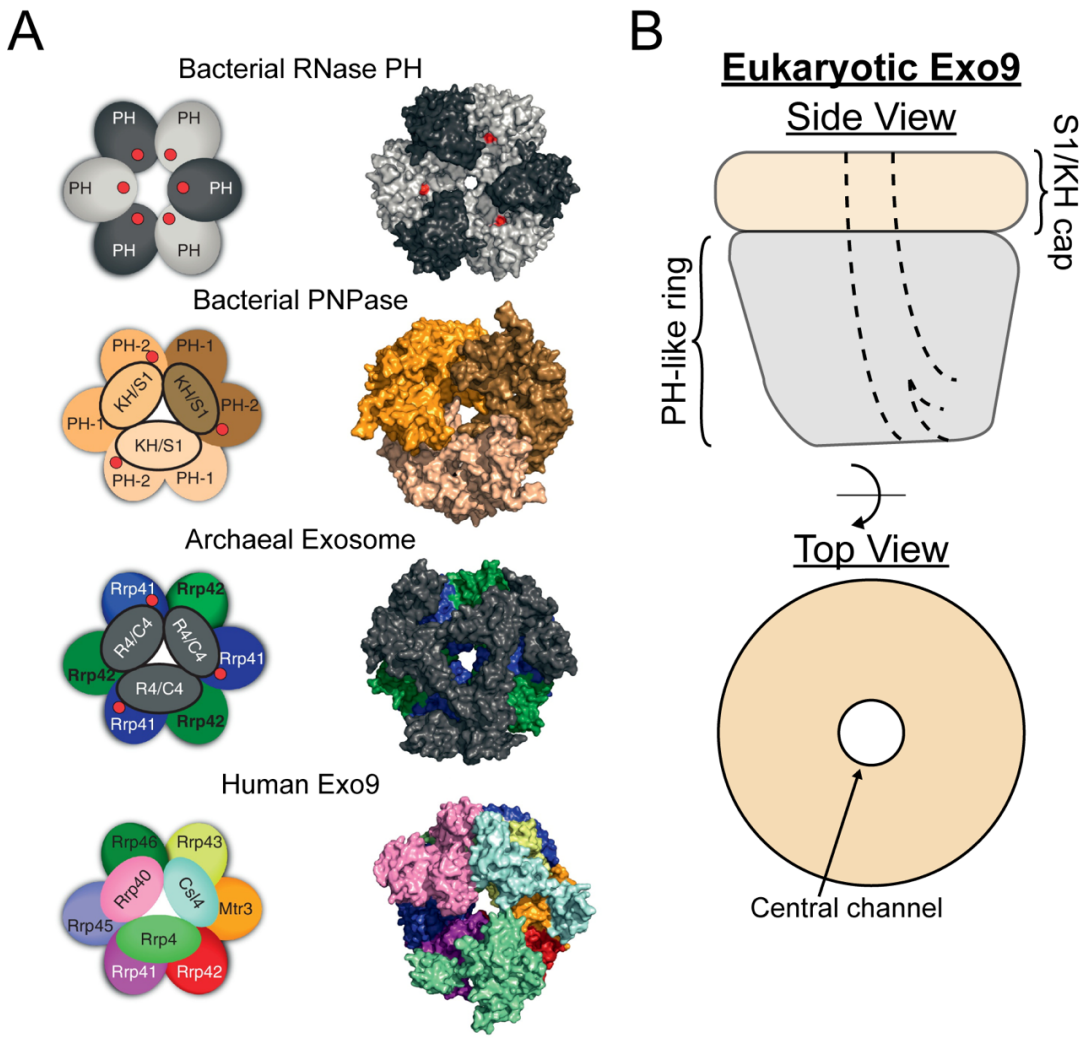


Figure 2. Conservation of architecture among 3' to 5' RNA degradation complexes. (A) Schematic (left) and surface representations (right) of top views of bacterial RNase PH (PDB 1UDN), bacterial PNPase (PDB 1E3P), the Rrp4 containing archaeal exosome (PDB 2JE6), and human Exo9 (PDB 2NN6). Phosphorolytic RNase active sites are depicted as red dots on the schematics. Modified from Figure 2 of Januszyk and Lima, 2014. (B) Simplified schematic representation of the side and top views of the eukaryotic Exo9 core. The central channel is represented by dotted lines in the side view.

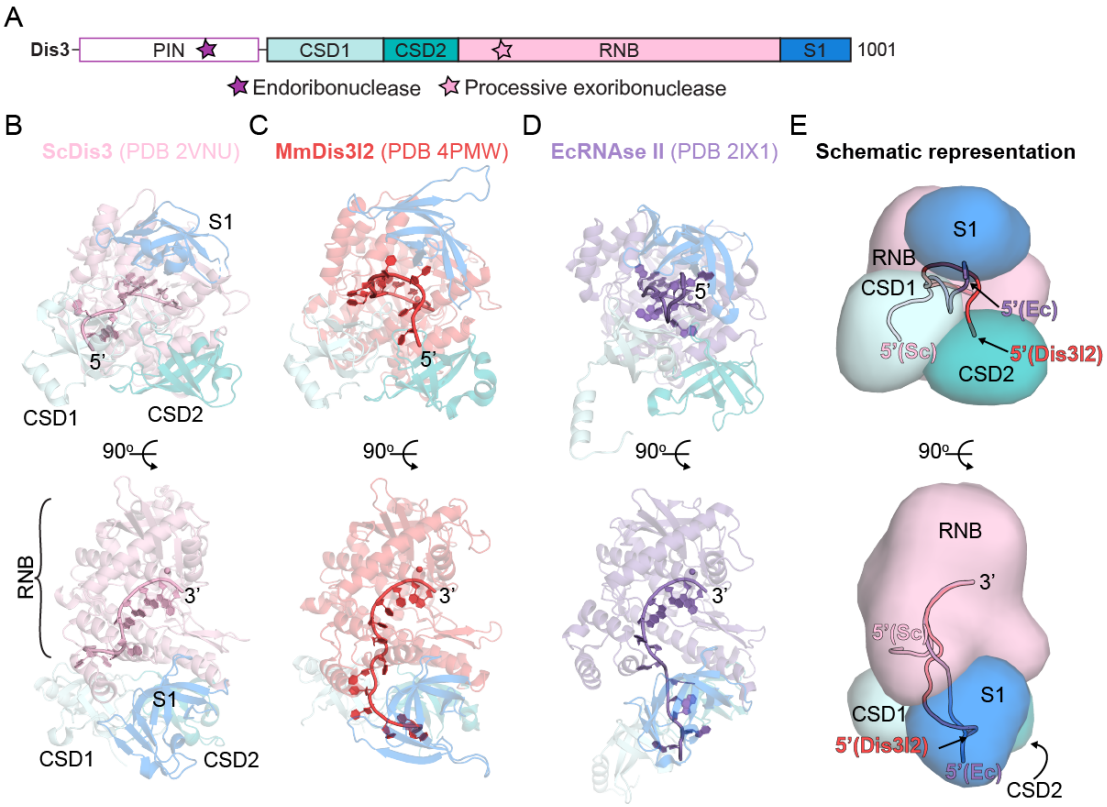


Figure 3. Comparison of domain architecture and RNA paths among Dis3 homologs. (A) Domain schematic for *S. cerevisiae* Dis3 (ScDis3). Active sites are highlighted with stars, and domains are colored to correspond with those in panels B through E. Amino acid length is shown. The PIN domain, which binds the exosome core and is not encompassed in the structure in panel B, is indicated with a white box. (B-D) RNA paths to the RNB active site of ScDis3 (B), mouse Dis3l2 (C) and the bacterial Dis3 homolog RNAse II (D). Structures were aligned to the RNB domain of ScDis3. E) Overlay of RNA paths to the RNB exonuclease site. Surfaces were generated in Pymol (Schrödinger) using domains from the ScDis3 structure (PDB 2VNU). RNA from all three structures is represented as a ribbon with 5' and 3' ends indicated.

the other two enzymes (Figure 3E). Dis3 from yeast and human include a second active site in the PilT N-terminal (PIN) domain that catalyzes distributive Zn^{2+}/Mn^{2+}-dependent endoribonuclease activity (Lebreton et al., 2008; Schaeffer et al., 2009; Schneider et al., 2008). The PIN domain also binds the PH-like ring subunits opposite to their surfaces that interact with the S1/KH cap (Bonneau et al., 2009), tethering Dis3 to the exosome. The nuclear subunit Rrp6, a homolog of bacterial RNAse D, includes a single active site that catalyzes Mg^{2+}-dependent distributive 3' to 5' hydrolytic exoribonuclease activity (Burkard and Butler, 2000; Targoff and Reichlin, 1985) (Figure 4). While Rrp6's C-terminal exosome interacting region (EAR) wraps around Exo9 making extensive contacts to Csl4, Mtr3, and Rrp43, the catalytic module, which includes the Exo domain, an N-terminal region and a helicase and RNase D C-terminal (HRDC) domain (Figure 4), rests atop the Exo9 on a conserved surface of cap proteins Rrp4 and Rrp40 (Makino et al., 2013b; Wasmuth et al., 2014) (Figures 1 and 5).

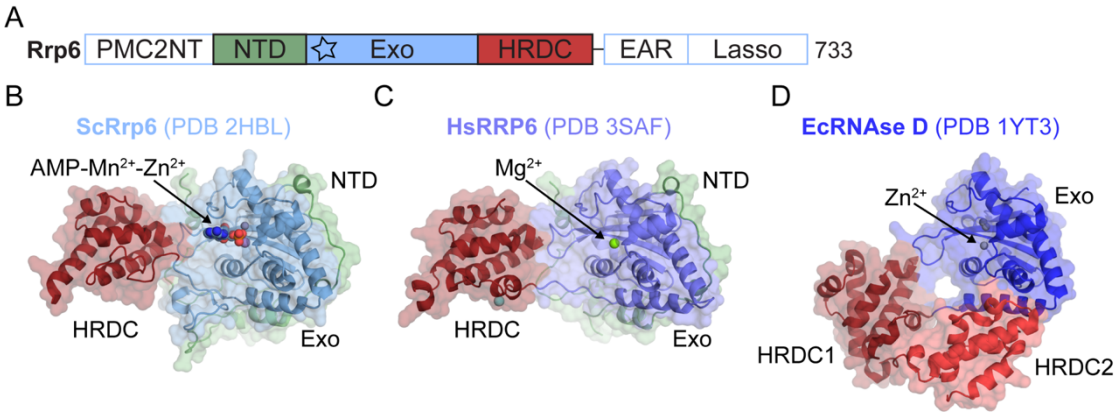


Figure 4. Structures of Rrp6 and RNase D. (A) Domain schematic for *S. cerevisiae* Rrp6 (ScRrp6). The active site is highlighted with a star, and domains are colored to correspond with those in panels B through D. Amino acid length is indicated. The PMC2NT, EAR, and lasso domains, which are not encompassed in the structure in panel B, are indicated with white boxes. (B-D) Crystal structures of the catalytic module of ScRrp6 (B), human Rrp6 (E) and the bacterial Rrp6 homolog RNAse D from *E. coli* (D). Structures were aligned based to the Exo domain of ScRrp6. AMP and/or metal ions proximal to the active site of each enzyme are highlighted.

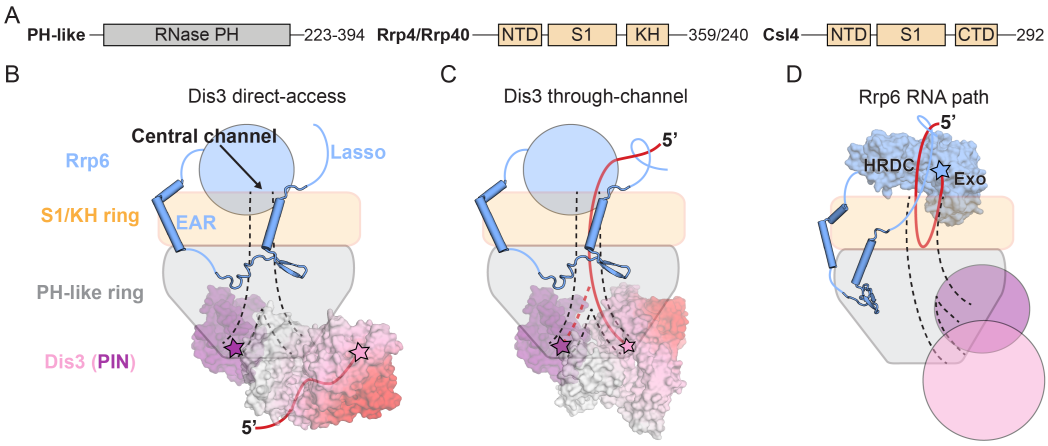


Figure 5. RNA paths and Dis3 conformations in the yeast nuclear exosome. (A) Domain schematics for *S. cerevisiae* Exo9 components. Amino acid lengths are indicated. (B) Direct-access conformation of Dis3. Dis3 and Rrp6 EAR domain are from PDB 5K36. The central channel is indicated by black dashed lines and RNA is represented as a red line with 5' end indicated. (C) Through-channel conformation of Dis3. Dis3 is from PDB 4IFD and Rrp6 EAR domain is from PDB 5K36. RNA and the central channel are indicated as previously, with the dashed red line representing a speculative RNA path to the Dis3 endo active site. (D) Rrp6 catalytic module (from PDB 5K36) bound to the core with RNA in its active site. RNA path to Rrp6 is based on biochemical data and PDB 5K36. Modified from Figure 1 of Zinder and Lima, 2017.

An early hypothesis posited that RNA degradation by the exosome was conceptually similar to protein degradation by the proteasome (van Hoof and Parker, 1999). In this model, the ribonuclease active sites of the complex are sequestered from the cellular milieu to prevent spurious degradation, and purposeful degradation requires that substrates be licensed to gain access to the active sites through a restricted channel. Subsequent studies largely confirmed this hypothesis (Makino et al., 2013a), including the observation that Rrp6 and Dis3 activities were modulated or inhibited when associated with the Exo9 core (Bonneau et al., 2009; Dziembowski et al., 2007; Wasmuth and Lima, 2012a), and that various surfaces within the Exo9 central channel were important for guiding RNA to the respective active sites in both yeast and human models (Drazkowska et al., 2013; Wasmuth and Lima, 2012a; 2017; Wasmuth et al., 2014; Zinder et al., 2016). Additionally, results from genetic studies in yeast showed that sterically occluding or mutating conserved basic residues lining the central channel led to substantial RNA processing and decay defects (Drazkowska et al., 2013; Wasmuth and Lima, 2012a). Subsequent structures also support these models: RNA can thread through the Exo9 central channel to reach Dis3 for processive degradation (Kowalinski et al., 2016; Liu et al., 2016; Makino et al., 2013b) or it can be deflected back to the Rrp6 active site for distributive processing or degradation (see below; Figure 5) (Wasmuth et al., 2014; Zinder et al., 2016).

Yeast Dis3 conformations and RNA paths

Structures of *S. cerevisiae* RNA exosome complexes with Dis3 revealed two prominent conformations for the enzyme (Bonneau et al., 2009; Liu et al., 2016; 2014; Makino et al., 2013b; 2015). While the PIN domain remains relatively static, the catalytic module rotates nearly 120 degrees between the two conformations (Figures 5B and 5C). One conformation can bind short ssRNAs (<14 nt), bypassing the Exo9

central channel (termed the direct-access or channel-independent Dis3 conformation, Figure 5B). This conformation features an extensive interaction surface with Exo9 and is also observed in the absence of RNA, suggesting that it is the resting state of the yeast RNA exosome (Bonneau et al., 2009; Liu et al., 2014; 2016). The other conformation is observed when Dis3 binds longer RNAs (>24 nt) that can pass through the Exo9 central channel (termed the through-channel or channel-dependent conformation, Figure 5C) (Liu et al., 2014; Makino et al., 2013b). This conformation features fewer interactions with Exo9, and is thought to be stabilized by the presence of long RNAs (Liu et al., 2016).

Rrp6 conformations and RNA paths

As previously mentioned, Rrp6 is tethered to the Exo9 core through a C-terminal EAR that wraps around the S1/KH cap and PH-like ring (Figures 4A and 5). In several structures, its catalytic domain is positioned atop the Exo9 core via interactions between the Rrp6 HRDC and Exo domains and a conserved surface on the S1/KH ring near the entrance to the central channel (Wasmuth et al., 2014; Makino et al., 2015; Zinder et al., 2016) (Figure 5). However, the catalytic module of Rrp6 can be displaced at equilibrium if a structured RNA with 3' overhang long enough to reach a catalytically inactivated Dis3 is present (Makino et al., 2015).

The model in which RNA accesses the Rrp6 catalytic site via interactions with the S1/KH ring proteins is supported by UV-crosslinking, biochemical analysis of complexes with mutant S1/KH or Rrp6 subunits (Wasmuth and Lima, 2012a; 2017), and two recent structures. These structures show the 3' end of RNA anchored to the Rrp6 active site, with the RNA path directed towards the S1/KH region of the central channel. While the remaining RNA was disordered in an earlier structure (Wasmuth et al., 2014), a more recent model showed that RNA can be deflected by the S1/KH ring

to position its 5' end near a channel formed between the HRDC and Exo domains of Rrp6 (Zinder et al., 2016) (Figure 5D and Chapter 1 of this dissertation).

While some models suggest that Rrp6 plays a passive role during Dis3-mediated RNA decay (Makino et al., 2015), other data suggest that Rrp6 can enhance Dis3 activities in the nuclear RNA exosome. One line of evidence supporting this is that degradation of poly(A)+ transcripts that accumulate in yeast strains lacking Rrp6 can be partially rescued by expressing catalytically inert Rrp6 (Assenholt et al., 2008; Mukherjee et al., 2016). Furthermore, in vitro studies showed that Rrp6 can activate the RNA decay activities of Dis3, especially evident for substrates with poly(A) tails (Wasmuth and Lima, 2012a; 2017; Wasmuth et al., 2014). Interestingly, Dis3 activation requires two portions of Rrp6, its catalytic module and its C-terminal tail, termed the RNA lasso (Wasmuth and Lima, 2017) (Figures 4A and 5B through 5D). While the catalytic module binds the S1/KH ring to presumably widen the central channel, the C-terminal domain binds RNA to enhance Rrp6 and Dis3 activities on a variety of RNA substrates. Although disordered in all available structures, it is perhaps noteworthy that the RNA lasso is positioned near the top of the Exo9 core where it could assist binding RNA adjacent to the central channel (Figure 5).

TRAMP and Ski complexes

The previous section focused on structure/activity relationships for the RNA exosome as a standalone machine, however it is likely that exosome cofactors mediate most encounters between RNA substrates and the RNA exosome. Recent studies focused on mechanisms that recruit these cofactors to the exosome and how their various activities influence RNA decay. We will restrict discussion in this section to the Trf/Air/Mtr4 polyadenylation (TRAMP) and Ski complexes, as their RNA helicase

components are important and conserved modulators of RNA exosome activities in the nucleus and cytoplasm, respectively. Intact structures of TRAMP or Ski complexes with the RNA exosome are not available, but structures of individual components or sub-complexes combined with genetic and biochemical studies support models as presented in Figures 6 and 7.

The TRAMP complex

The TRAMP complex was initially uncovered by analysis of a temperature sensitive mutant of *S. cerevisiae* that expressed hypomodified $tRNA_i^{Met}$ whose phenotype was suppressed by mutations in a non-canonical nuclear poly(A) polymerase (Kadaba et al., 2004). This polymerase (Trf4) and its paralog (Trf5) were found to exist in complexes with zinc knuckle proteins Air1 or Air2, and the DExH helicase Mtr4 (LaCava et al., 2005; Vanacova et al., 2005; Wyers et al., 2005) (Figure 6). TRAMP is thought to assist nuclear degradation in yeast, including decay of cryptic unstable transcripts (CUTs) and processing of rRNAs (LaCava et al., 2005; Vanacova et al., 2005; Wyers et al., 2005). The TRAMP complex genetically interacts with the RNA exosome to promote RNA degradation via its 3' non-templated poly(A) polymerase and RNA helicase activities, presumably via physical interactions between the helicase and exosome (Chen et al., 2001; Falk et al., 2017; Schuch et al., 2014; Wasmuth et al., 2017). While Mtr4 is essential and likely integrated into other complexes, an *mtr4* allele that lacks ATP binding activity fails to rescue lethality of a Δ*mtr4* strain (Taylor et al., 2014). With respect to TRAMP, Air1/Air2 are dispensable for viability (LaCava et al., 2005), however simultaneous deletion of *trf4* and *trf5* results in lethality (Castaño et al., 1996). These observations underscore the

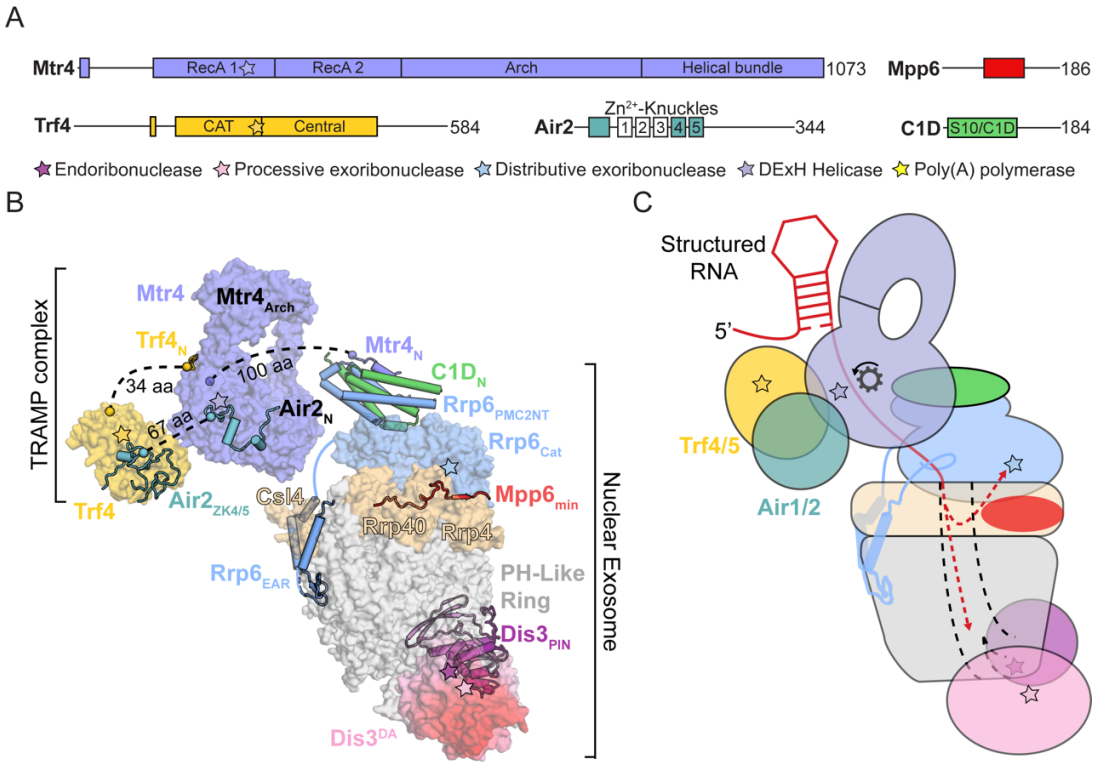


Figure 6. The TRAMP complex and the nuclear exosome. (A) Domain schematics for *S. cerevisiae* C1D, Mpp6, and TRAMP components. Catalytic sites are indicated as stars and amino acid lengths are shown. (B) Structural models for the nuclear exosome and associated cofactors with RNA omitted for clarity. Black dotted lines represent connecting regions for which no structural information is available. Mtr4 and Trf4/Air2 peptides are from PDB 4U4C; Trf4/Air2 zinc-knuckles are from PDB 3NYB; PH-like ring, Rrp40, Rrp4, Csl4, Dis3, Rrp6 catalytic module, and Rrp6 EAR are from PDB 5K36; minimal Mpp6 ($Mpp6_{min}$) is from 5VZJ; Rrp6 PMC2NT domain, C1D, and Mtr4 N-terminal peptide are from PDB 4WFD and were positioned based on PDB 5C0W. (C) Model for Mtr4 threading of RNA to the nuclear exosome after polyadenylation by Trf4/5. The central channel is indicated by black dashed lines and RNA is represented as a red line with 5' end indicated. Dashed red arrows represent RNA paths to the catalytic subunits. Helicase direction is indicated by a gear and arrow. Modified from Figure 2 of Zinder and Lima, 2017

importance of a functioning TRAMP complex for nuclear RNA exosome function and for viability.

Crystal structures of Mtr4 revealed a multi-domain helicase core that resembles the bacterial DNA repair helicase Hel308, with a flexible insertion termed the arch domain that is specific to Mtr4 and Ski2 helicases (Jackson et al., 2010; Weir et al., 2010) (Figure 6). Biochemical and structural studies showed that Trf and Air proteins form a stable heterodimer that can be isolated from tagged Mtr4 using a high-salt wash, and its interactions with the helicase core of Mtr4 occur via short peptide motifs (Falk et al., 2014; Hamill et al., 2010; LaCava et al., 2005; Losh et al., 2015) (Figure 6B). It is presumed that the polyadenylation activities of TRAMP are important for generating 3' single stranded tails that are long enough to be captured by Mtr4 for unwinding (Figure 6C) or by Rrp6 or Dis3 for deadenylation or degradation, respectively.

Nuclear cofactors that bridge Mtr4 and the exosome

C1D is a small protein with functions in RNA metabolism (Mitchell, 2010) and the DNA damage response (Jackson et al., 2016). While often referred to as a nuclear exosome cofactor, the observations that C1D is critical for stable Rrp6 expression and it is present in approximately stoichiometric amounts in endogenous nuclear complexes support its inclusion as a primary subunit of the nuclear exosome (Feigenbutz et al., 2013a; Shi et al., 2015). GST-C1D from humans can pulldown in vitro translated Rrp6 and knockdown of Rrp6 shifts eGFP-C1D from being concentrated in nucleoli to being evenly distributed throughout the nucleus, supporting a similar role for C1D in higher eukaryotes (Schilders et al., 2007). A structure of the nuclear exosome bound to C1D revealed interactions with the N-terminal PMC2NT domain of Rrp6 and its position above the Rrp6 catalytic module, forming a 'lid'

above the exosome (Makino et al., 2015). The composite interface between Rrp6 and C1D binds a small peptide motif near the N-terminus of Mtr4, providing a physical tether to the exosome (Schuch et al., 2014) (Figure 6), while C-terminal region of C1D binds nucleic acid in vitro and interacts with protein components of Box C/D snoRNPs (Costello et al., 2011). With that said, the N-terminal domain of C1D fully rescues growth in synthetic lethal Δ*rex1*Δ*rrp47* and Δ*mpp6*Δ*rrp47 S. cerevisiae* strains, suggesting that the most critical functions for C1D may pertain to Mtr4/TRAMP recruitment and stabilization of Rrp6 (Costello et al., 2011; Feigenbutz et al., 2013a; Garland et al., 2013).

Mpp6 is another small, nucleic acid binding protein that associates with the Exo9 core (Falk et al., 2017; Schilders, 2005; Schuch et al., 2014; Wasmuth et al., 2017). Cross-linking experiments coupled to mass spectrometry showed that it contacts the S1/KH protein Rrp40 in the yeast complex (Shi et al., 2015), and competition binding experiments at high concentrations reported low affinity interactions between its C-terminal region and the Rrp6 Exo domain (Kim et al., 2016). Subsequent structural and biochemical studies of yeast proteins corroborated the mass spectrometry data (Figure 6B) and further showed that a small conserved region in Mpp6's N-terminus can recruit Mtr4 to the exosome (Falk et al., 2017; Wasmuth et al., 2017) (see Chapter 2). This recruitment enables the exosome to degrade a dsRNA with a short 3' overhang and to cover a longer (~50 nt vs ~30 nt) fragment of RNA in RNase protection assays (Falk et al., 2017; Wasmuth et al., 2017). Human Mpp6 also interacts with Mtr4 (Chen et al., 2001; Schilders et al., 2007) and the exosome core (Falk et al., 2017), supporting a role for Mpp6 in recruitment of Mtr4 or Mtr4 complexes in higher eukaryotes.

The Ski complex

The cytoplasmic Ski complex consists of a DExH box helicase Ski2, a tetratricopeptide repeat scaffold protein Ski3, and two copies of the beta propeller protein Ski8 (Fig. 7) (Brown et al., 2000; Halbach et al., 2013; van Hoof et al., 2000). The Ski complex contributes to mRNA turnover, degradation of aberrant mRNAs, viral defense (Molleston and Cherry, 2017), and RNAi pathways in some eukaryotes (Łabno et al., 2016a). Deletion of any subunit of the Ski complex in *S. cerevisiae* results in synthetic lethality when combined mutations of decapping enzymes or deletion of the 5'-3' exoribonuclease Xrn1 (Anderson and Parker, 1998; Araki et al., 2001; Johnson and Kolodner, 1995; van Hoof et al., 2000). Ski2 contains an N-terminal region that is necessary for Ski complex interactions followed by a helicase core and a flexible insertion domain, similar to features observed in Mtr4 (Halbach et al., 2013; 2012; Wang et al., 2005) (Fig. 7). Ski2 is the only member of the Ski complex with a clear homolog in humans (SKIV2L), however it plays important roles in viral defense and elimination of Rig-I-like receptor ligands generated by Ire1 during the unfolded protein response, presumably through association with the human exosome (Aly et al., 2016; Eckard et al., 2014).

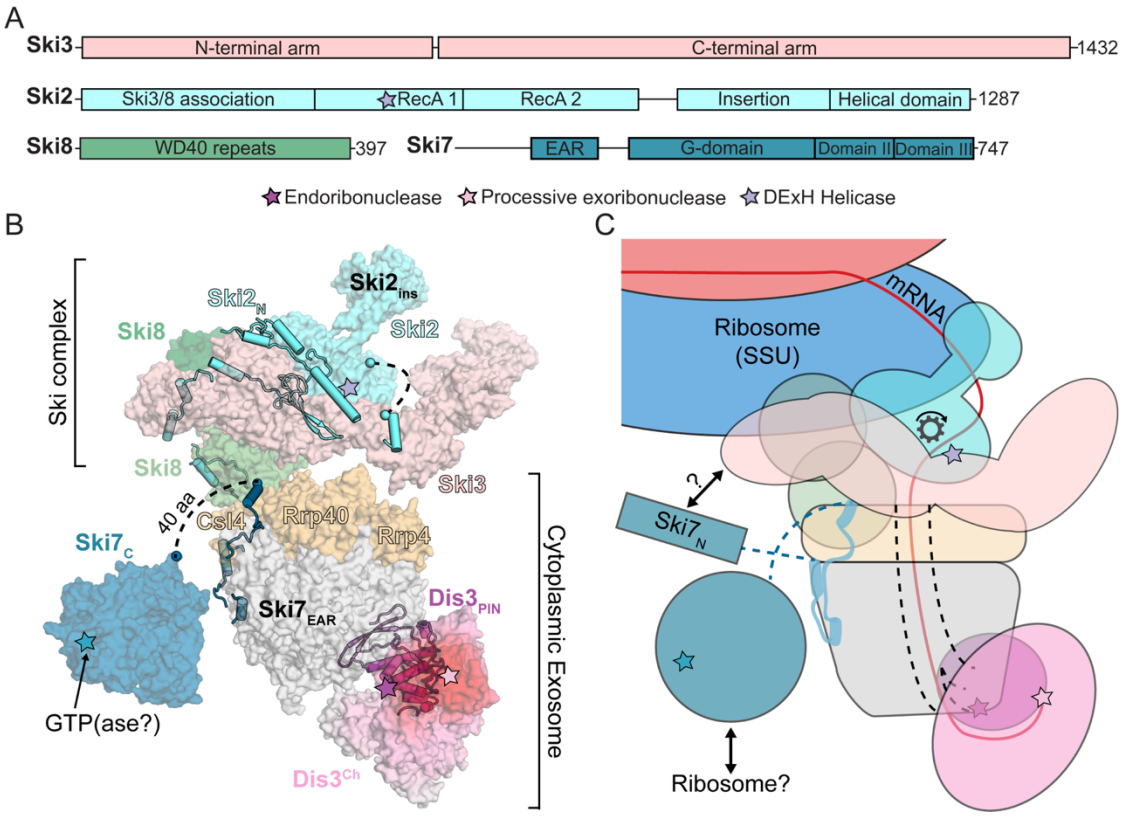


Figure 7. The Ski complex and the cytoplasmic exosome. (A) Domain schematics for *S. cerevisiae* Ski complex components. Catalytic sites are represented as stars and amino acid lengths are shown. (B) Structural models for the cytoplasmic exosome and associated cofactors with RNA omitted for clarity. Black dotted lines represent connecting regions for which no structural information is available. Ski3, Ski8 and Ski2 N-terminus are from PDB 4BUJ; Ski2 globular region and insertion are from PDB 4A4Z and were aligned to PDB 4BUJ; Ski7 C-terminal domains are from PDB 4ZKE; PH-like ring, Rrp40, Rrp4, Csl4, Dis3, and Ski7 EAR are from PDB 5JEA. (C) Model for Ski complex channeling of RNA to the cytoplasmic exosome. The central channel is indicated by black dashed lines and RNA is represented as a red line with 3' end shown bound to the Dis3 exonuclease active site. Helicase direction is indicated by a gear and arrow. Figure 3 of Zinder and Lima, 2017.

RNase protection assays further suggest that association of the Ski complex with the exosome extends the through-channel RNA path to Dis3 by ~10 nt, leading to a model for channeling (Halbach et al., 2013) (Figure 7C). Channeling through the helicase of the Ski complex was directly observed in a recent cryo-EM structure of the Ski complex bound to a ribosome containing an mRNA plus a 3' overhang (Schmidt et al., 2016). In this structure, the small subunit (SSU) of the ribosome contacts the Ski2 helicase, the N-terminal TPR region of Ski3, and one of the Ski8 proteins, and the mRNA overhang is seen channeled through Ski2's pore. Notably, Ski2's arch insertion inhibits RNA binding in the context of the Ski complex bound to the exosome (Halbach et al., 2013), but binding of the arch to the ribosome stabilizes it in a conformation that enables easy accessibility of the RNA 3' end to the helicase core (Schmidt et al., 2016).

Ski7/HBS1Lv3 bridges the Ski complex and RNA exosome

Yeast Ski7 contains a globular C-terminal GTPase domain and N-terminal domains that bridge the RNA exosome and Ski complexes (Araki et al., 2001; van Hoof et al., 2000; Wang et al., 2005). A crystal structure of the C-terminal domain of *S. cerevisiae* Ski7 revealed its structural similarity to active translational GTPases, though GTPase activity could not be confirmed in vitro (Kowalinski et al., 2015). Aligned 2D class averages from negative stain EM of endogenous *S. cerevisiae* Ski7 containing cytoplasmic exosome complexes suggests that the C-terminal globular domains of Ski7 adopts multiple conformations when bound to the exosome, and 3D reconstructions from cryo-EM revealed that the Ski7 N-terminal domain interacts with the Exo9 core via surfaces that overlap with those used by the Rrp6 C-terminal domain (Liu et al., 2016) (Figures 6 and 7). This latter result was confirmed with binding assays and observed in a contemporaneous crystal structure (Kowalinski et al.,

2016). Additionally, two groups independently identified a short splicing isoform of human HBS1L (HBS1Lv3) as the long sought-after Ski7 homolog and confirmed its ability to interact with the human exosome and Ski complex (Kalisiak et al., 2016; Kowalinski et al., 2016). The canonical HBS1L isoform in humans contains a C-terminal GTPase fold that isoform 3 lacks, but it does not interact with the exosome, suggesting that multiple and unique sub-complexes of the Ski complex may exist in higher eukaryotes.

Targeting RNAs to the exosome and associated complexes

The RNA exosome can cooperate with its cofactors to specifically target transcripts for degradation. Selective targeting has purported roles in diverse processes such as rRNA processing, transcription-coupled decay, suppression of untimely meiosis in *S. pombe*, as well as viral defense (Aly et al., 2016; Harigaya et al., 2006; Kilchert et al., 2016; Lee et al., 2013; Molleston and Cherry, 2017; Molleston et al., 2016). RNAs can be targeted for decay based on recognition of specific sequences or structural elements by adaptor proteins and subsequent recruitment of the exosome or associated cofactor complexes. Additionally, the integrity of the RNA, which can be compromised via misfolding, RNA polymerase errors, or lack of proper post-transcriptional modifications, may be sensed by the exosome or associated cofactors and specifically target that RNA for decay over its correctly generated counterparts.

Suppression of untimely meiosis in S. pombe

Perhaps the most extensively studied system for selective targeting of RNAs based on sequence elements by the exosome is that of meiotic transcript degradation in *S. pombe*. Under normal growth conditions, meiotic RNAs such as *mei4* are

transcribed but rapidly targeted for degradation, thus preventing cells from unintentionally entering meiosis (Harigaya et al., 2006). The regions responsible for degradation of these transcripts were found to contain multiple repeats of a U(U/C)AAAC sequence that acted as determinants of selective removal (DSR) (Harigaya et al., 2006; Yamashita et al., 2012). When these sequences were appended to reporter genes that would normally be expressed at high levels, they were strongly repressed except during meiosis (Harigaya et al., 2006). It was found that DSR sequences in *S. pombe* are specifically recognized by the YTH domain of a protein called Mmi1, which in turn recruits the RNA exosome via an interaction with the MTREC complex (discussed below) (Egan et al., 2014; Harigaya et al., 2006; Lee et al., 2013; Wang et al., 2016). During nitrogen starvation, Mei2, a master regulator of meiosis in *S. pombe*, sequesters Mmi1 into a single nuclear focus, allowing its RNA targets to accumulate and meiosis to proceed (Harigaya et al., 2006). Recently, DSRs were found in introns of non-meiotic *S. pombe* genes and were shown to regulate expression of those RNAs in an intron-retention dependent manner (Kilchert et al., 2015). In cases where co-transcriptional splicing fails to remove these introns, the resulting transcript is recognized by Mmi1 and degraded by the exosome.

Cofactor mediated bridging to the cap-binding complex

In addition to TRAMP, human Mtr4 is present in at least two other complexes, the nuclear exosome targeting (NEXT) and poly(A) tail exosome targeting (PAXT) complexes (Lubas et al., 2011; Meola et al., 2016; Ogami et al., 2017). These complexes include mutually exclusive interactions between Mtr4, a zinc finger protein (Zcchc8 in NEXT and ZFC3H1 in PAXT), and an RNA binding protein (Rbm7 in NEXT and Pabpn1 in PAXT) (Figure 1B). The NEXT complex promotes degradation of PROMPTs and 3’ extended RNAs in the nucleoplasm (Lubas et al., 2011; Preker et

al., 2008). PAXT promotes degradation of longer and more mature (with respect to poly(A) tail length) substrates compared to NEXT substrates (Meola et al., 2016; Ogami et al., 2017). Both NEXT and PAXT interact with the cap-binding complex containing Ars2 (CBCA complex) via an adaptor protein ZC3H18, physically tethering the exosome to nascent capped transcripts to promote degradation following termination (Andersen et al., 2013; Meola et al., 2016). In *S. pombe*, a nuclear Mtr4-like helicase Mtl1 associates with a zinc finger protein Red1 and as many as four other proteins to form the MTREC complex, which is involved in maintenance of facultative heterochromatin and degradation of meiotic RNAs, CUTs, and unspliced transcripts (Egan et al., 2014; Kilchert et al., 2015; Lee et al., 2013; Zhou et al., 2015). Because it also interacts with the CBCA complex and the exosome and Red1 is homologous to ZFC3H1, MTREC is possibly a fission yeast counterpart to the human PAXT complex. Importantly, NEXT, PAXT, and MTREC complexes lack 3' polyadenylation activity, suggesting that their RNA binding and helicase activities are sufficient to generate 3' ssRNA that is long enough to engage the exosome.

Mtr4/AIM interactions for selective RNA decay

Ribosomal RNAs are derived from two transcripts in *S. cerevisiae*, one that codes for the 18S, 5.8S, and 25S rRNAs and the other codes for the 5S rRNA (Figures 8A and 8B) (Henras et al., 2014). A set of proteins known as the processome dynamically associates with the pre-rRNA co-transcriptionally to direct endo- and exonucleolytic RNA processing in addition to chaperoning ribosomal proteins within the RNP (Woolford and Baserga, 2013).

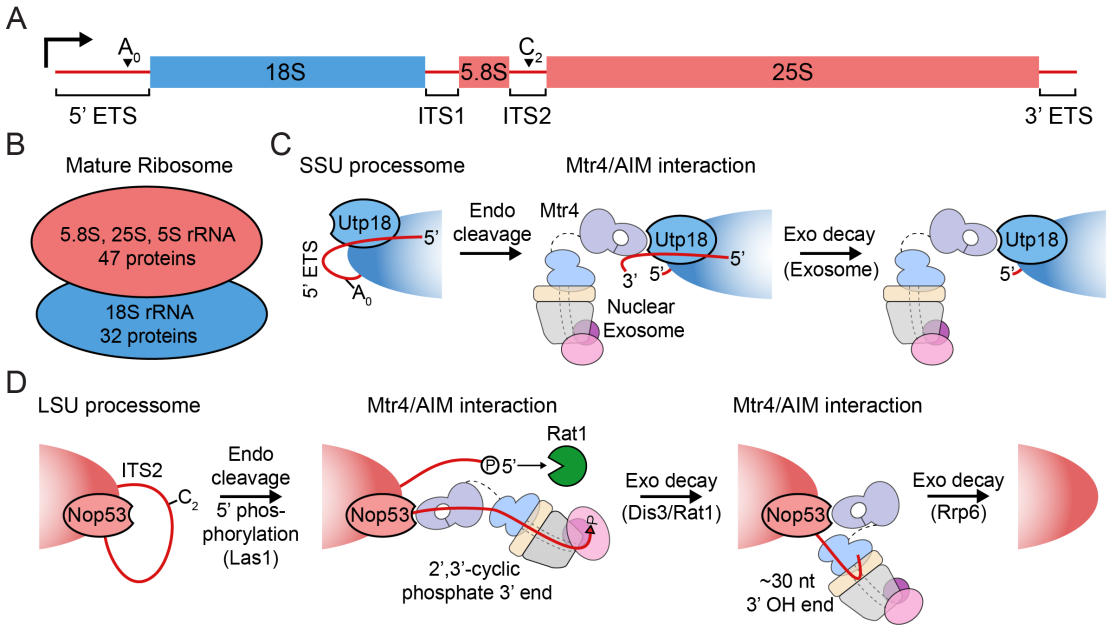


Figure 8. AIM-arch interactions recruit the exosome for rRNA processing (A) Schematic for the *S. cerevisiae* 7 kb pre-rRNA molecule. Regions of the RNA contained within the mature ribosome are shown as boxes and spacers are shown as lines. Direction of transcription is shown with an arrow and A_0 and C_2 endonucleolytic cleavage sites are indicated. (B) Schematic for a mature ribosome. (C) A $Utp18_{AIM}$ - $Mtr4_{Arch}$ interaction recruits the exosome for 5' ETS removal after endonucleolytic cleavage at the A_0 site. (D) A $Nop53_{AIM}$ - $Mtr4_{Arch}$ interaction recruits the exosome for 5.8S rRNA processing after LasI cleavage at the C_2 site. Modified from figure 4 of Zinder and Lima, 2017.

A recent study examining rRNA processing uncovered a conserved motif (arch interacting motif or AIM) in processome proteins Utp18 and Nop53 that interacts with the arch of Mtr4 to recruit the exosome (Thoms et al., 2015). They found that interaction of the Mtr4 arch with this motif in Utp18, a subunit of the early-associating SSU processome subcomplex UtpB, enables removal of the 5' externally transcribed spacer (5' ETS) while interaction with Nop53, an LSU processome factor, enables 3' processing of the 7S precursor RNA (Figures 8C and 8D). Mutation of the AIM motif on either Nop53 or Utp18 results in accumulation of unprocessed precursors for their respective substrates (Thoms et al., 2015). Interestingly, other proteins such as Air2 and Sqs1 can interact with the Mtr4 arch in a similar region as Nop53/Utp18, suggesting that hierarchical competition for the arch may contribute to RNA decay in the nucleus (Losh and van Hoof, 2015).

A recent high-resolution cryo-EM structure of the *S. cerevisiae* SSU processome positions the N-terminus of Utp18 proximal to the 3'-most nucleotide that could be modeled from the 5' ETS (Barandun et al., 2017). Neither the Utp18 AIM, which resides near the its N-terminus, nor the 3' end of the 5' ETS that would be expected from A_0 cleavage could be modeled, though their approximate locations are close enough to be compatible with recruitment of the Mtr4-exosome to this site for degradation (Barandun et al., 2017).

3' end chemistry and RNA fate

Dis3 or Rrp6 mediated degradation of model substrates by exosomes containing Dis3 and Rrp6 appears stochastic in vitro, however this seems an unlikely strategy for RNA decay in vivo. While nuclear cofactor complexes can influence RNA fate as discussed, recent studies suggested that the chemical structure of the 3' end might influence the fate of a particular RNA. For example, Rrp6 from humans and *S.*

cerevisiae lacks activity on synthetic RNAs containing a 3’ phosphate, while Dis3 family enzymes readily degrade these RNAs (Burkard and Butler, 2000; Domanski et al., 2016; Lubas et al., 2013; Tomecki et al., 2010; Zinder et al., 2016). While RNAs with 3’ phosphate may not be abundant in vivo, cleavage by exo- and endoribonucleases generates 2’,3’-cyclic phosphate ends for tRNA introns, U6 snRNA, and some ribosomal RNA precursors (Gasse et al., 2015; Knapp et al., 1979; Lund and Dahlberg, 1992). Structural models suggest that Dis3 but not Rrp6 could degrade RNAs with 2’,3’-cyclic phosphate (Lorentzen et al., 2008; Zinder et al., 2016).

. For U6 snRNA, a component of the spliceosome, Mpn1/Usb1 generates a 2’,3’-cyclic phosphate at the 3’ end as a product of its exonuclease activity, and U6 RNA is polyadenylated and rapidly degraded in Δ*mpn1 S. pombe* strain (Shchepachev et al., 2012; 2015). This suggests that a 2’,3’-cyclic phosphate at the 3’ end stabilizes the U6 RNA by preventing decay by Rrp6 or polyadenylation by TRAMP. Indeed, a transcriptome-wide survey of 2’,3’-cyclic phosphate RNAs revealed U6 is by far the most abundant RNA with this chemical signature in HeLa cells, though this approach may have missed other RNAs that are targeted for rapid degradation (Schutz et al., 2010). Interestingly, Usb1 from *S. cerevisiae* was recently demonstrated to possess cyclic phosphodiesterase activity that cleaves the terminal 2’,3’-cyclic phosphate to generate a 3’ phosphate, which it is unable to further trim into shorter products (Didychuk et al., 2017). This additional activity, which the human enzyme lacks, may prevent ScUsb1 from over-shortening the U6 3’ poly(U) tail, which the authors of the study note would not be a concern for organisms that could extend over-shortened tails via TUTases (Didychuk et al., 2017).

To generate the 7S rRNA (5.8S rRNA with a 3’ extension), the Las1 endonuclease component of the tetrameric Las1 complex cleaves the rRNA precursor

molecule at the C_2 site of internally transcribed spacer 2 (ITS2) (Gasse et al., 2015) (Figures 8A and 8D), resulting in a 2',3'-cyclic phosphate at the 3' end of the 7S rRNA and a 5'OH on the other fragment. While the latter RNA is 5' phosphorylated by the Grc3 kinase subunit of the complex to enable processing by Rat1/Rai1 (the two other components of the Las1 complex), the 2',3'-cyclic phosphate end is an intermediate prior to processive 3' to 5' processing by Dis3 in the exosome (Figure 8D). The produces the 5.8S rRNA plus ~30 nt at the 3' end, the approximate length required to span the Exo9 central channel. This overhang is subsequently removed by Rrp6 (Allmang et al., 1999a; Briggs et al., 1998) leading to the handoff model presented in Figure 7D. Interestingly, the length of 5.8S overhang produced in siRNA knockdowns of human Rrp6 is 40 nt, suggesting a longer through-channel RNA path for the human exosome (Tafforeau et al., 2013). While this and other 2',3'-cyclic phosphate 3' ends may be resolved by phosphodiesterases prior to processing or degradation, it is likely that 3' modified RNAs could be substrates of Dis3 and the nuclear RNA exosome. Supporting this, a recent study reconstituted 5.8S processing using purified components from *S. cerevisiae* (Fromm et al., 2017). By placing a tag on Nop53, the investigators were able to purify a population of LSU processomes containing a significant amount of 7S precursor RNA. They incubated these particles with WT or mutant exosome complexes, Mtr4, and ATP in different combinations and observed rRNA processing via northern blotting and negative stain EM. Their experiments showed that Dis3's endo and exo activities, Mtr4's helicase activity, and ATP were all required for efficient removal of the 7S product. Consistent with in vivo data, inactivation of Rrp6 minimally affected the disappearance of the 7S species, which presumably contains a 2',3'-cyclic phosphate at its 3' end, but resulted in the accumulation of a 5.8S + 30 nt species (Fromm et al., 2017).

The RNA exosome and its roles in cellular homeostasis

Because the Exo9 components and Dis3 are essential for viability in single-celled organisms, it is perhaps unsurprising that the exosome contributes to important and diverse biological processes in higher eukaryotes and is mutated in several diseases (Fabre and Badens, 2014; Robinson et al., 2015; Staals and Pruijn, 2010). Here we present several recent advances in our understanding of how the exosome and its cofactors contribute to proliferation, differentiation, innate immunity against RNA viruses, and telomerase RNA surveillance.

Proliferation and differentiation

Dis3 has gained notoriety for its role in cellular proliferation, and was identified as one of the most frequently mutated genes in genome-wide association studies of multiple myeloma (MM) (Chapman et al., 2011; Lohr et al., 2014; Walker et al., 2012). Most mutations observed in these studies cluster within its exoribonuclease domain and are predicted to disrupt its 3' to 5' decay activity. As Dis3 activities are generally associated with promoting cell division (Ohkura et al., 1988; Tomecki et al., 2014), inactivation of Dis3 in MM was somewhat perplexing. A recent study addressed this by characterizing inactivating mutations in Dis3 in *Drosophila*, *C. elegans*, and mouse models (Snee et al., 2016). While mutating Dis3 alone resulted in mitotic defects, increased RAS activities acted synergistically with this mutant to stimulate growth, a phenotype that was not evident using activated RAS alone. This perhaps explains the observation that KRAS activities were often increased in MM clones that contained inactivating DIS3 mutations (Lohr et al., 2014). Furthermore, another study depleted Dis3 in human MM and other cell-lines

and observed accumulation of Let28b, a protein that sequesters the *let-7* family of micro RNAs to prevent their maturation (Segalla et al., 2015). Because *let-7* RNAs can silence *MYC*, *RAS*, and other mRNAs, Dis3 depletion ultimately results in accumulation of these gene products, potentially explaining correlations observed for DIS3 inactivation and RAS activation in model systems (Snee et al., 2016) and MM (Lohr et al., 2014). Perhaps consistent with this model, decreased Dis3 expression has been observed in high risk genotypes associated with pancreatic cancer (Hoskins et al., 2016), where activating RAS mutants are common (Eser et al., 2014).

Recent work has also illuminated a role for the exosome during erythropoiesis, the process through which hematopoietic stem cells differentiate into erythrocytes (McIver et al., 2014; 2016). In this process, the balance between hematopoietic stem cell differentiation and proliferation is critical: too much proliferation can lead to tumor formation while excessive differentiation can exhaust the supply of stem cells. For erythropoiesis, proliferation and terminal differentiation are enforced by stem cell factor (SCF) and erythropoietin, respectively. GATA-1 and Foxo-3 are master transcription factors that control differentiation during erythropoiesis, and both down-regulate expression of EXOSC8, an Exo9 subunit (Table 1). Interestingly, shRNA knockdown of exosome core subunits in hematopoietic stem cells resulted in an accumulation of GATA-1 and Foxo-3 regulated transcripts, suggesting that the exosome may counter differentiation by degrading these transcripts in the absence of erythropoietin (McIver et al., 2014), similar to a role proposed for the exosome in maintaining a proliferative state in human skin stem cells via selective targeting of the *GRLH* mRNA (Mistry et al., 2012). Furthermore, hematopoietic stem cells depleted of exosome components were non-responsive to SCF due to decreased levels of its cognate receptor tyrosine kinase, Kit, though they remained responsive to erythropoietin (McIver et al., 2016).

Viral defense

A role for the exosome and its cofactors in viral defense was described nearly two decades before its discovery through a genetic screen that identified the '*SKI*' genes in *S. cerevisiae*. The *SKI* genes were so named because of the "superkiller" phenotype observed: mutations in *SKI* genes increased levels of a killer toxin that was produced by the M viral dsRNA (Ridley et al., 1984; Toh-E et al., 1978). It was later discovered that three of these proteins (Ski2, Ski3, and Ski8) form the Ski complex (Brown et al., 2000) that interacts with the RNA exosome via Ski7 (Araki et al., 2001; van Hoof et al., 2000) (Figures 1A and 7). Other *SKI* genes were later identified as subunits of the exosome itself (Table 1).

A more recent study using cultured human cells revealed a role for the Ski complex in antiviral defense against Hepatitus B virus (HBV) (Aly et al., 2016). A screen for helicases that could suppress HBV replication identified the human Ski complex RNA helicase SKIV2L (Table 1). It was further demonstrated that interactions between the HBV X-RNA, SKIV2L, HBS1L (HsSki7) and the exosome resulted in selective degradation of the HBV X-RNA via the NSD pathway.

The metazoan TRAMP complex was recently shown to participate in viral defense in the cytoplasm (Molleston et al., 2016). Infection of human and *Drosophila* cells with the disparate RNA viruses vesicular stomatitis virus, Sindbis virus, or Rift Valley fever virus (RVFV) was potentiated by knockdown of exosome and TRAMP components. While normally restricted to the nucleus and concentrated in the nucleolus (Lubas et al., 2011), TRAMP subunits are exported to the cytoplasm during infection where they participate in the degradation of viral RNAs (Molleston et al., 2016). Furthermore, appending a 3' UTR from RVFV to a GFP reporter was sufficient to stimulate its degradation upon RVFV infection in human cells, suggesting that viral

RNA sequences are targeted for selective degradation under these conditions. Mechanisms underlying TRAMP export to the cytoplasm and targeting of viral 3' UTRs await further investigation.

Telomerase RNA quality control

Several recent studies implicated the exosome and its cofactors in degradation and quality control of telomerase RNA (hTR) in HeLa cells (Nguyen et al., 2015a; Shukla et al., 2016; Tseng et al., 2015). Degradation of hTR is stimulated by the 3' polyadenylation activity of the human TRAMP complex and antagonized by the poly(A) binding protein Pabpn1 and the deadenylase PARN, which is mutated in some cases of the premature aging disease dyskeratosis congenita (DKC). One study further showed that knockdown of nuclear RNA decay machinery could rescue telomerase RNA levels and defects in telomerase activity in cells depleted of dyskerin, a protein subunit of the telomerase RNP that is also mutated in DKC, prompting the authors to suggest that the exosome could be a therapeutic target for certain telomere pathologies (Shukla et al., 2016).

DGCR8, a dsRNA binding protein involved in miRNA biogenesis, has been recently implicated as an adaptor protein for exosome targeting to structured substrates such as hTR (Macias et al., 2015). DGCR8 contains dsRNA binding and heme domains, which interact with the stem and apical regions of the pri-miRNA, respectively, as a dimer (Nguyen et al., 2015b). In the nucleoplasm, this dimer interacts with DROSHA, an RNA endonuclease involved in miRNA maturation, to ensure its fidelity in producing miRNAs (Macias et al., 2013). Investigators found that DGCR8 also interacts in a distinct complex with the nucleolar exosome, and this interaction is necessary for the turnover of snoRNAs and hTR in that compartment (Macias et al., 2015).

Outstanding questions

Contained in this dissertation are our attempts to answer several outstanding questions regarding the structure and functions of the exosome. It had long been appreciated that RNA exosome complexes are recalcitrant towards crystallization except in the presence of specific RNA substrates that presumably stabilize the enzymatic subunits in discrete conformations. One persistent challenge had been to structurally characterize RNA exosome complexes containing both Rrp6 and Dis3, as each enzyme requires an RNA 3' end and both engage RNA via channeling through the Exo9 pore. Chapter 1 describes a strategy we successfully employed to overcome this difficulty, namely synthesizing an RNA with two 3' ends (3'-3' RNA), and co-crystalizing the complex with it. Subsequent characterization of mutants designed based on the structure are described as well as a biochemical investigation of the degradation of a 3' phosphate RNA by the nuclear exosome.

Other questions remained that related to the biochemical functions of protein cofactors of the nuclear RNA exosome. In Chapter 2, we report a crystal structure of the same complex bound to a similar 3'-3' RNA and a minimal exosome-associating region of the nuclear cofactor Mpp6. Biochemical investigation that followed showed that Mpp6 and Rrp47 cofactors both contribute to recruitment of Mtr4 for helicase-dependent RNA degradation using purified *S. cerevisiae* proteins. Finally, Chapter 3 outlines our unpublished efforts towards determination of high-resolution structural models for Mtr4-dependent RNA degradation using cryo-EM

Author's note

All experiments contained in this dissertation were performed by the author except where explicitly noted in the figure legend. Experiments performed by individuals other than the author were included because they were essential to provide context for the author's work and for discussion in the text. Much of the content has been previously published in either Zinder and Lima, 2017 (Introduction), Zinder et al., 2016 (Chapter 1), and Wasmuth et al., 2017 (Chapter 2).

CHAPTER 1: Crystal Structure of the Nuclear Exosome

Introduction

The RNA exosome, a conserved multi-protein complex, processes and degrades RNA in the 3' to 5' direction in eukaryotes (Houseley and Tollervey, 2009; Januszyk and Lima, 2014). Substrates of the nuclear RNA exosome include Cryptic Unstable Transcripts (CUTs) (Wyers et al., 2005), PROMoter uPstream Transcripts (PROMPTs) in humans (Preker et al., 2008), hypomodified tRNAs (Kadaba et al., 2004), sn(o)RNAs (Allmang et al., 1999a), and ribosomal RNAs (Mitchell et al., 1997). Mutations that result in loss of function of the RNA exosome have been identified in hematopoietic malignancies (Chapman et al., 2011) and neurodegenerative disorders (Fabre and Badens, 2014), underscoring its physiological importance.

The RNA exosome core (Exo9) includes nine subunits, a six-membered ring formed by the RNase PH-like proteins Rrp45, Rrp41, Rrp43, Rrp46, Rrp42, and Mtr3 (the PH-like ring), that is capped by a three-membered ring of S1 and KH domain containing proteins Rrp4, Rrp40, and Csl4 (the S1/KH ring). The structure of Exo9 from human revealed a donut-shaped architecture with a prominent central channel large enough to accommodate single-stranded RNA (Liu et al., 2006) (Figure 2). Both yeast and human Exo9 cores are devoid of catalytic activity (Dziembowski et al., 2007; Liu et al., 2006).

In the budding yeast *S cerevisiae*, the cytoplasmic RNA exosome includes the Exo9 core and the processive exoribonuclease Rrp44 (aka Dis3) to form a ten-component complex ($Exo10^{44}$). In the nucleus, $Exo10^{44}$ associates with the distributive exoribonuclease Rrp6 (Figure 1A) to form an eleven-component complex ($Exo11^{44/6}$).

Protein cofactors such as Rrp47, Mpp6, and the TRAMP complex associate with the nuclear exosome to target particular RNA substrates or to alter its activities (Houseley and Tollervey, 2009; Januszyk and Lima, 2014; LaCava et al., 2005; Lubas et al., 2012; Wasmuth and Lima, 2012b; Wyers et al., 2005). All protein subunits of $Exo10^{44}$ are essential for cell viability in yeast (Brouwer et al., 2001; Dziembowski et al., 2007; Mitchell et al., 1997), while deletion of Rrp6 results in temperature sensitivity, slow-growth, and RNA processing defects (Allmang et al., 1999a; 1999b; Briggs et al., 1998).

Previous efforts to understand how Exo9 contributes to the RNA decay activities of the exosome revealed that the central channel was essential, that Rrp6 and Rrp44 activities are attenuated or altered when bound to the RNA exosome core, that RNA passes through the PH-like and S1/KH rings to access the Rrp44 active site, that RNA engages the S1/KH proteins to access the Rrp6 active site, and that RNA path(s) to the catalytic subunits are partially overlapping (Bonneau et al., 2009; Drazkowska et al., 2013; Makino et al., 2013b; Malet et al., 2010; Wasmuth and Lima, 2012a; Wasmuth et al., 2014).

A crystal structure of $Exo10^{44}$ bound to a fragment of Rrp6 and a stem-loop RNA with a 3' $polyU_{30}$ single-stranded extension at 2.8 Å (PDB 4IFD) showed RNA passing through the Exo9 central channel and a conformation for Rrp44 that differed from that observed in the absence of RNA (Bonneau et al., 2009; Makino et al., 2013b). A similar conformation of Rrp44 was observed in a recent 2.7 Å structure of $Exo10^{44}$ bound to an N-terminal fragment of the cytoplasmic RNA Exosome cofactor Ski7 and RNA (Kowalinski et al., 2016). Negative stain EM studies on $Exo10^{44}$ at 20-25 Å resolution (Liu et al., 2014) revealed two conformations for Rrp44, one in the absence of RNA or in the presence of short RNAs or structured RNAs with short 3' overhangs (termed the direct-access or channel-independent Rrp44 conformation) and

the other conformer observed in the presence of RNAs with single-stranded 3' ends long enough to span the Exo9 central channel (termed the through-channel or channel-dependent Rrp44 conformation). When $Exo10^{44}$ was degrading a structured RNA that included a long single-stranded 3' extension, sub-classification of EM data suggested that both conformations co-exist. More recent cryo-EM structures captured these same two conformations for $Exo10^{44}$ in complex with Ski7, now at 4.2 Å for the channel-independent Rrp44 conformation and 5.8 Å for the channel-dependent Rrp44 conformation (Liu et al., 2016).

A structure of Exo9 bound to Rrp6 ($Exo10^{6exo\text{-}\Delta N\Delta C}$) and 24 nt poly(A) RNA at 3.35 Å resolution (PDB 4OO1) revealed an RNA path to Rrp6 that partially overlapped with the channel-dependent RNA $Exo10^{44}$ path (Wasmuth et al., 2014), confirming predictions based on biochemical and genetic methods (Wasmuth and Lima, 2012a). Based on mutational analysis, a similar path was posited for the human $Exo10^{6}$ complex (Wasmuth et al., 2014). A twinned crystal structure reported at 4.6 Å resolution (69.5% overall completeness) of the nuclear exosome bound to the cofactor Rrp47 and 18 nt AU-rich RNA (PDB 5C0W) revealed its overall architecture, a speculative path for RNA bypassing Rrp6, and RNA bound to the channel-independent conformation of Rrp44 (Makino et al., 2015). Atomic resolution features of the nuclear exosome containing both Rrp6 and Rrp44, including the RNA path to Rrp6 and elements that stabilize the channel-independent conformation of Rrp44, remain unclear due to the nominal resolution of structures to date.

Here we report a crystal structure at 3.1 Å resolution of a nuclear $Exo11^{44/6\Delta N\Delta C}$ complex with RNA bound to both Rrp6 and Rrp44 active sites. The structure reveals extensive RNA contacts to Rrp6 and the S1/KH subunits, Rrp44 in a channel-independent conformation bound to RNA, elements of the Exo9 core that interact with Rrp44 in the channel-independent conformation and inhibit its activities,

and details within the Rrp44-RNA active site that provide a structural basis for its ability to degrade 3' phosphorylated RNA substrates.

Results

3'-3' RNA and the nuclear RNA exosome structure

High-resolution structures of the nuclear exosome have not been achieved thus far, possibly because the Rrp44 and Rrp6 catalytic subunits are dynamic when not bound to RNA substrates. As both catalytic subunits require a 3' end to bind RNA in their respective active sites, we employed Cu^{I}-catalyzed alkyne-azide cycloaddition, or 'copper click chemistry,' to synthesize an RNA with two 3' ends that we hypothesized would enable simultaneous capture of both catalytic subunits (Figures 9A and 10A through 10D). After surveying different RNA lengths, we obtained crystals that diffracted to 3.1 Å resolution of the *S. cerevisiae* nuclear exosome ($Exo11^{44exo\text{-}endo\text{-}/6exo\text{-}\Delta N\Delta C}$) in complex with an RNA composed of two 17 nt segments linked by 19 carbon to carbon, nitrogen, or oxygen bonds (Figure 9). This complex contains all full-length, catalytically inactive proteins except Rrp6, which lacks its C-terminal RNA-binding 'lasso' portion (Wasmuth and Lima, 2017) and N-terminal Rrp47 interaction domain (Figure 4A) (Stead et al., 2007). A molecular replacement solution was obtained using *S. cerevisiae* $Exo10^{6exo\text{-}\Delta N\Delta C}$ (PDB 4OO1) and Rrp44 from the Rrp44-Rrp41-Rrp45 complex (PDB 2WP8) as search models (Bonneau et al., 2009; Wasmuth et al., 2014), and iterative rounds of building and refinement resulted in an Rwork/Rfree of 0.201/0.249 with good stereochemistry for a structure that includes 3731 amino acids and 24 nucleotides of RNA (Table 2).

Figure 9. Architecture of Exo11$^{44/6}$ bound to 3'-3' RNA. (A) Domain schematics of Rrp6 and Rrp44. The shaded area indicates the amino acid boundaries used for crystallization. Active site aspartates that were mutated to asparagine for binding assays and crystallization are indicated. (B) Schematic for synthesis of a 34 nt RNA molecule with two 3' ends. RNA is represented as an arrow pointing from 5' to 3'. Gel is 15% acrylamide TBE-urea and RNA is visualized using SYBR gold stain. The lane for the reaction containing 2:1 [RNA]:[diazide] is indicated with an asterisk. (C) Surface representation of Exo11$^{44/6}$ bound to 3'-3' 34 nt RNA. RNA is shown as a cartoon with 3' and 5' ends of the model indicated. (D) RNA path in Exo11$^{44/6}$/3'-3' RNA crystals. Mesh is a simulated annealing 2Fo-Fc omit map contoured at 1.0 σ. The dashed lines indicate the most likely path connecting 5' ends of the RNA fragments.

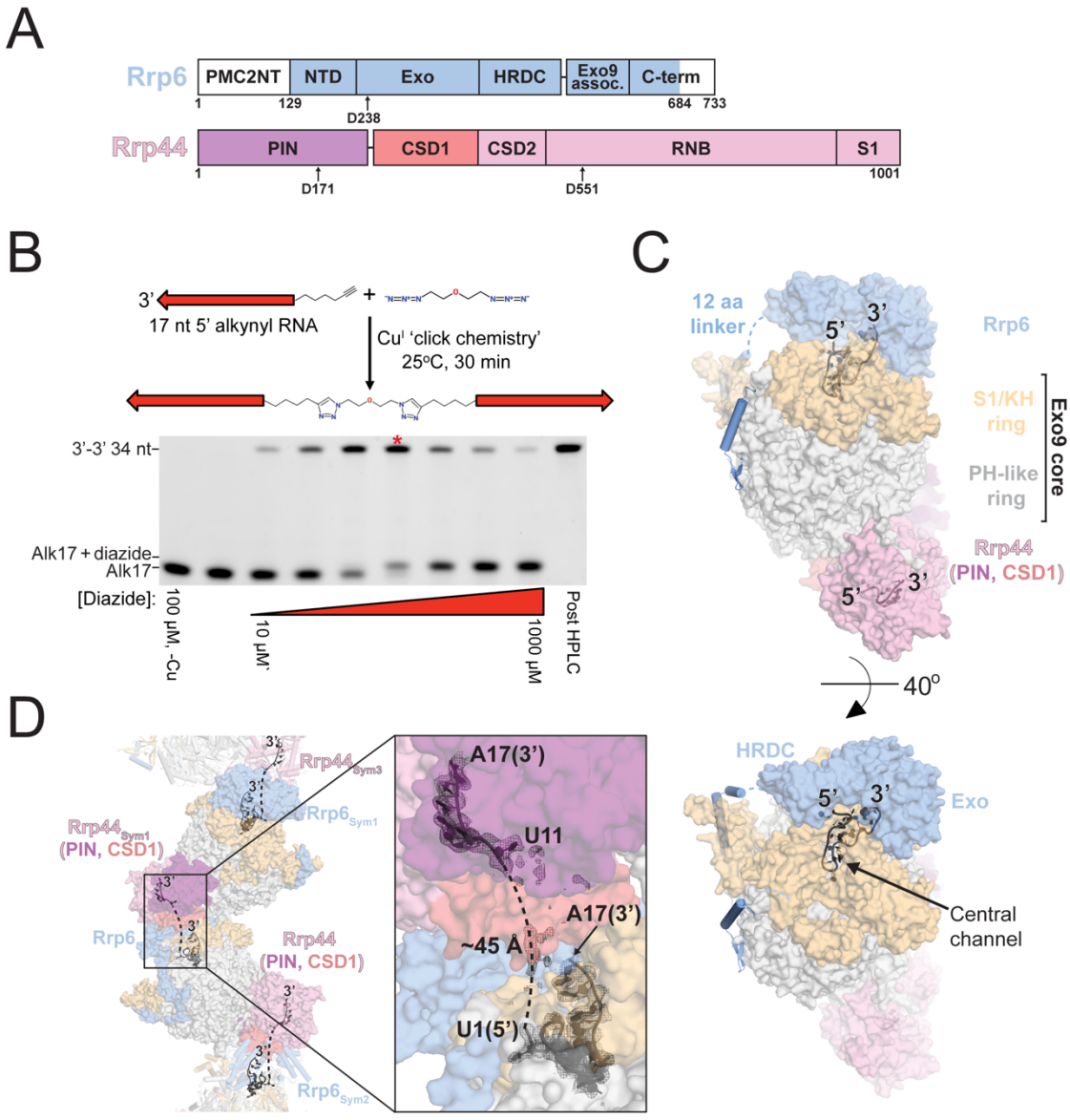

A
Rrp6
PMC2NT
NTD
Exo
HRDC
Exo9 assoc.
C-term
1
129
D238
684
733
Rrp44
PIN
CSD1
CSD2
RNB
S1
1
D171
D551
1001
B
3'
17 nt 5' alkynyl RNA
Cu[I] 'click chemistry'
25°C, 30 min
3'-3' 34 nt
Alk17 + diazide
Alk17
[Diazide]:
100 µM, -Cu
10 µM
1000 µM
Post HPLC
C
12 aa linker
5'
3'
Rrp6
S1/KH ring
PH-like ring
Exo9 core
Rrp44 (PIN, CSD1)
40°
HRDC
Exo
Central channel
D
Rrp44Sym3
Rrp6Sym1
Rrp44Sym1 (PIN, CSD1)
Rrp6
Rrp44 (PIN, CSD1)
Rrp6Sym2
A17(3')
U11
~45 Å
U1(5')

Figure 10. Purification and characterization of 3'-3' RNA and Exo11$^{44/6}$/3'-3' RNA crystals. (A) Chromatograms from DEAE separation of an RNA click reaction using an NaCl gradient. Buffers A and B are as described in the Materials and Methods section. For this experiment, 135 µg of total RNA was loaded in a volume of 500 µL. B) 15% polyacrylamide TBE-urea gel of fractions, which were taken every minute from the HPLC run in (A). RNA is visualized using SYBR gold stain. C and D) Rrp44 decay time courses on 5'-3'(C) or 3'-3' (D) 36 nt AU-rich RNA substrates. Rrp44 is at a concentration of 5 nM and substrate RNA is at a concentration of 50 nM. Decay intermediates apparent between the alkynyl 25 and alkynyl 18 marker in panel C indicate that Rrp44 cannot degrade through the di-triazole linker. RNA is visualized using SYBR gold staining. E and F) ~10 crystals were transferred from the mother liquor into three separate 2 µl drops of well solution (washes) and then dissolved in 20 µL of 20 mM Tris-Cl pH 8.0, 100 mM NaCl 1 mM TCEP-HCl. Volumes in the 'Crystal' lane indicate how much of this solution was loaded in that lane. For the protein gel (E), 10µL of dissolved crystal solution was added to 5 µl of 4x loading dye and 5 µL of ultrapure water and 10 µL of that solution was run on a 4-12% acrylamide BIS-TRIS run in MOPS-SDS buffer and stained with SYPRO Ruby. Standards are 1.3, 0.65 and 0.32 pmol of purified complex. For the RNA gel (F), 5 µL crystal solution was added to 5 µL 10 mM Tris-Cl pH 8.0, 0.2 % w/v SDS, 10 mM EDTA, 0.1 U/µL proteinase K and incubated at 50°C for 10 minutes. Following this incubation, 10 µL of 2x urea-PAGE loading buffer was added and 10 µL of that was run on a 15% acrylamide TBE-urea gel. Standards are 1.2, 0.59, and 0.29 pmol of RNA. Quantification of Rrp44 and RNA bands yields an approximate protein to RNA molar ratio of 1:1.3.

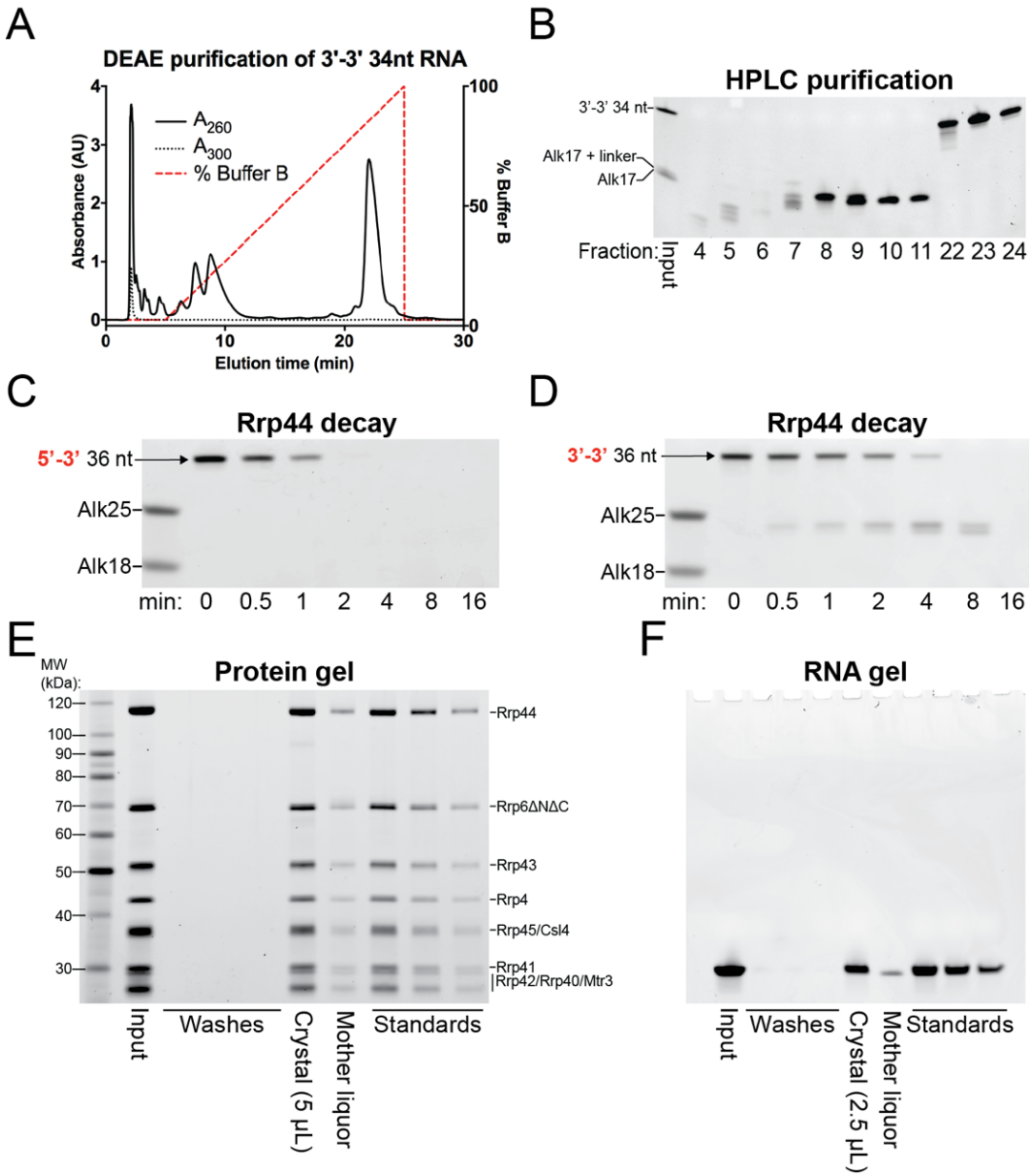

A
DEAE purification of 3'-3' 34nt RNA
A260
A300
% Buffer B
Absorbance (AU)
% Buffer B
Elution time (min)
B
HPLC purification
3'-3' 34 nt
Alk17 + linker
Alk17
Fraction: Input 4 5 6 7 8 9 10 11 22 23 24
C
Rrp44 decay
5'-3' 36 nt
Alk25
Alk18
min: 0 0.5 1 2 4 8 16
D
Rrp44 decay
3'-3' 36 nt
Alk25
Alk18
min: 0 0.5 1 2 4 8 16
E
Protein gel
MW (kDa):
120
100
90
80
70
60
50
40
30
Rrp44
Rrp6ΔNΔC
Rrp43
Rrp4
Rrp45/Csl4
Rrp41
Rrp42/Rrp40/Mtr3
Input
Washes
Crystal (5 µL)
Mother liquor
Standards
F
RNA gel
Input
Washes
Crystal (2.5 µL)
Mother liquor
Standards

Table 2. Crystallographic data and refinement statistics

	$Exo11^{Rrp44/Rrp6/RNA}$
Data collection	
X-ray Source	APS NE-CAT 24IDC
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	140.7, 212.3, 218.2
α, β, γ (°)	90.0, 90.0, 90.0
Wavelength (Å)	0.9795
Resolution (Å)	106-3.1 (3.21-3.1) *
R_{merge}	0.111 (0.534)
$I/\sigma I$	14.3 (2.8)
$CC_{1/2}$	0.994 (0.495)
Completeness (%)	98.8 (97.0)
Redundancy	3.6 (3.0)
Wilson B factor ($Å^2$)	66.2
Refinement	
Resolution (Å)	106-3.1
No. reflections observed	423414
No. unique reflections	106037
R_{work}/ R_{free}	0.201/0.249
No. atoms	30252
Protein	29377
RNA	502
Ligands	92
Water	280
Average B-factors	
Protein	80
RNA	107
Ligands	96
Water	52
R.m.s deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.43
Ramachandran plot	
% favored	94.6
% allowed	5.2
% outliers	0.16
Molprobity	
Clashscore/Percentile	$5.7/100^{th}$
MolProbity Score/Percentile	$1.68/100^{th}$

One crystal was used. Highest resolution shell is shown in parenthesis.

The overall architecture of the complex (Figure 9C) reveals Rrp6 and Rrp44 at opposite ends of the Exo9 central channel with Rrp6 bound to RNA above the S1/KH ring and Rrp44 below the PH-like ring bound to RNA in the channel-independent Rrp44 conformation. This conformation of Rrp44 resembles that observed in previous structures (Bonneau et al., 2009; Liu et al., 2014; 2016; Makino et al., 2015), while the Rrp6 conformation resembles that observed for $Exo10^{6exo\text{-}\Delta N\Delta C}$ bound to poly(A) (Wasmuth et al., 2014), a position that is rotated away from the S1/KH ring in a model of the nuclear exosome structure with Rrp6 bound to the cofactor Rrp47 and not bound to RNA (Makino et al., 2015). Distance constraints and electron densities suggest that the 3'-3' RNA traverses between Rrp6 in one Exo11 complex to Rrp44 in an adjacent complex (symmetry mate) rather than traveling through the central channel of a single complex (Figure 9D).

Contacts between the Exo9 core and Rrp44 contribute to Rrp44 activities

Previous structural studies observed a channel-independent conformation of Rrp44 in association with the Exo9 core (Bonneau et al., 2009; Liu et al., 2014; 2016; Makino et al., 2015), but none were determined at sufficient resolution to reveal all of the features of the Exo9 core proteins that stabilize this conformation. We now observe that Rrp43 residues 100-120 (Rrp43L1) form a loop followed by a pair of helices that contact Rrp44 near the point of RNA ingress (Figures 11A, 11B, and 12). Inspection of electron density in a previous structure revealed the presence of Rrp43L1 however it was not modeled (Makino et al., 2015) (Figure 13). In addition, the C-terminus of Rrp45 ($Rrp45_{Cterm}$) forms an extended strand that contacts Rrp44 near the 3' end of its bound RNA in our structure as previously observed (Bonneau et al., 2009; Liu et al., 2016; Makino et al., 2015).

Figure 11. Features in the non-catalytic Exo9 core contact the RNA channel-independent and channel-dependent conformations of Rrp44. (A) Illustration of the channel-independent and channel-dependent conformations of Rrp44 along with associated features in the Exo9 core from the side and top views. The center panel depicts a theoretical transition between the two states in which the Rrp43L1, Rrp43L2 and $Rrp45_{Cterm}$ are disordered. Surface representations of Rrp44 were generated using the model from this work and PDB 4IFD for the channel-independent and channel-dependent conformations, respectively. (B and C) Rrp43L1, Rrp43L2 and $Rrp45_{Cterm}$ features associated with the channel-independent (B) and channel-dependent (C) conformations of Rrp44 from this work and PDB 4IFD, respectively. Order/disorder boundaries of Rrp43L1 and Rrp43L2 are indicated by colored spheres.

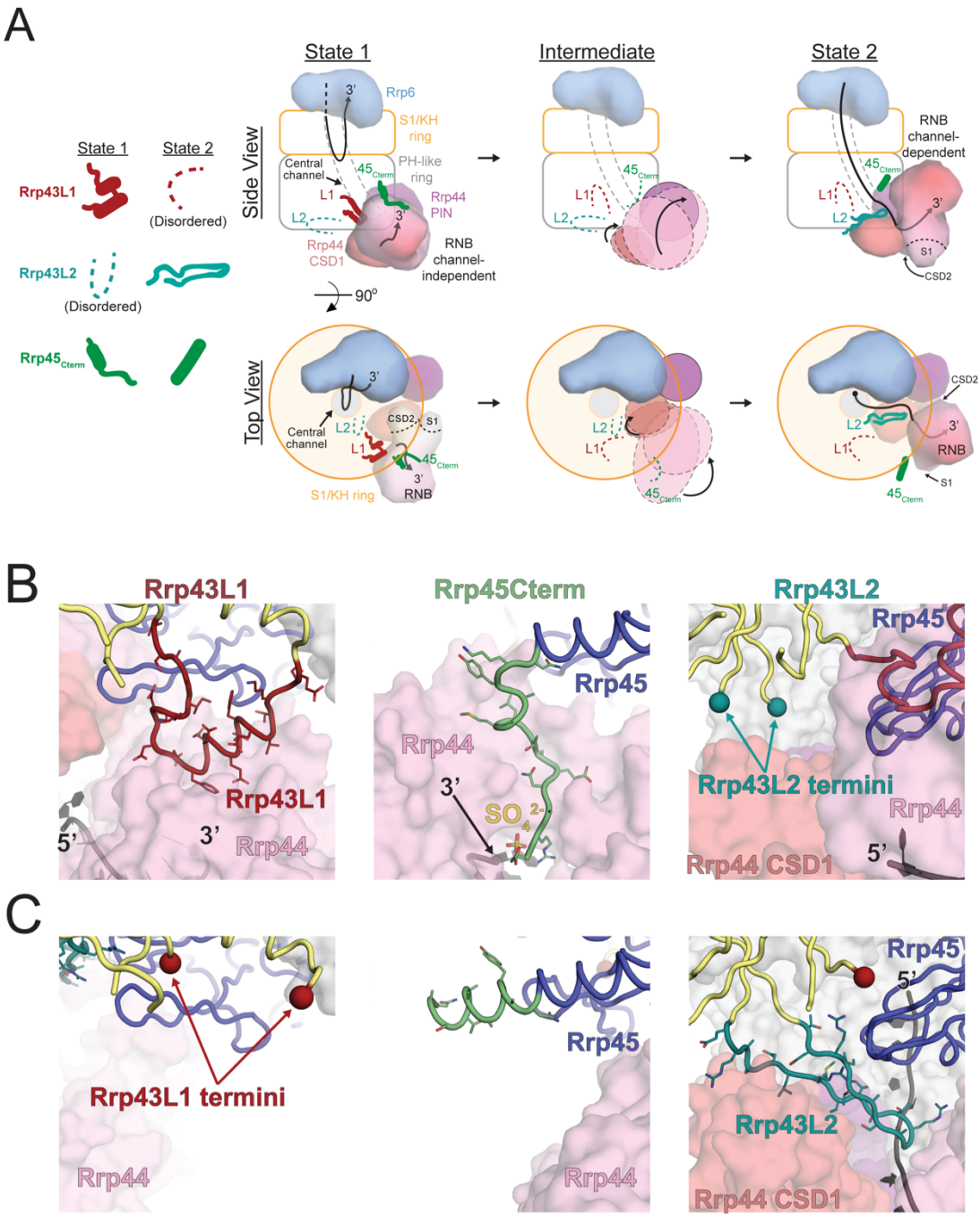
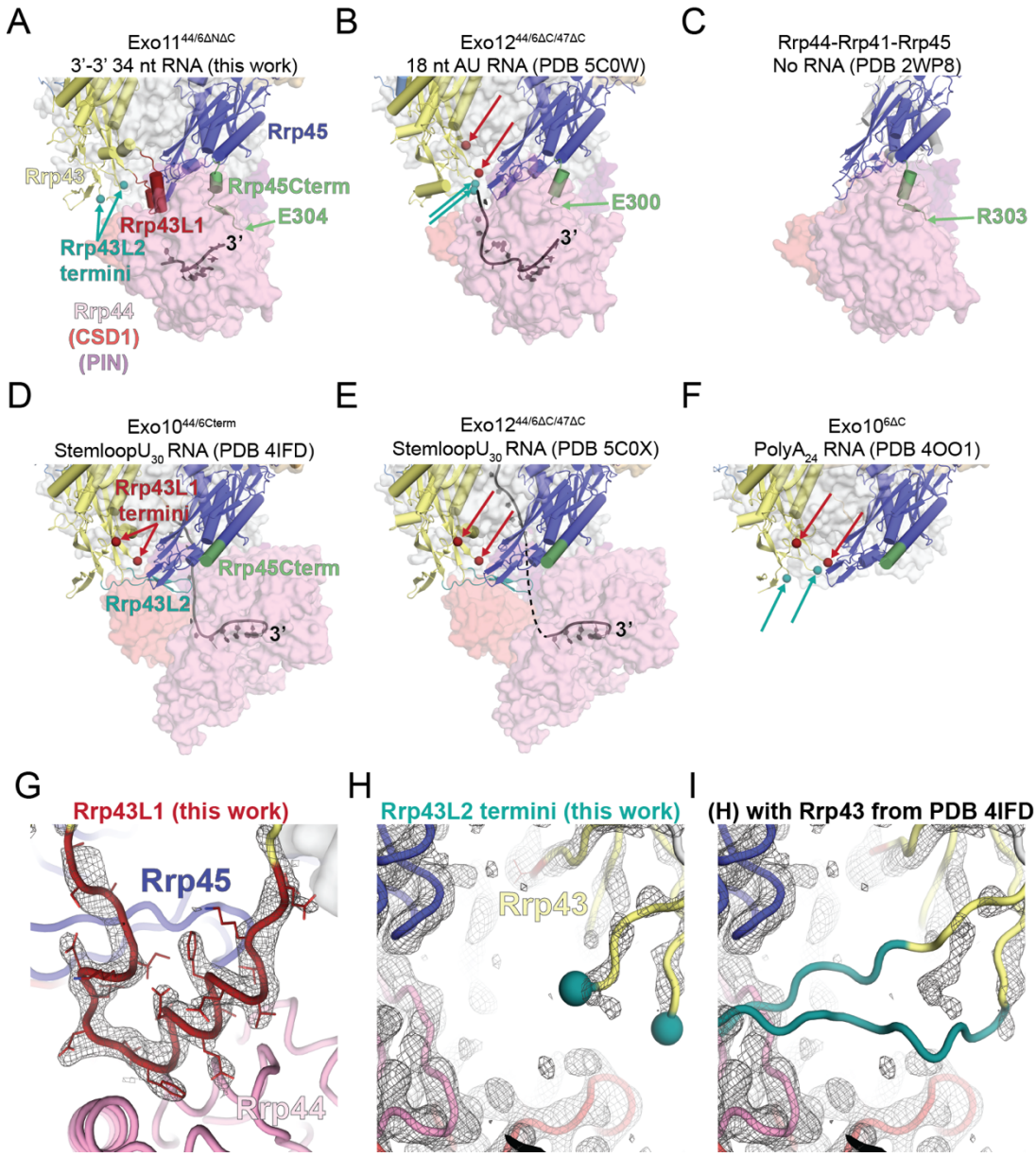

A
State 1
State 2
Rrp43L1
(Disordered)
Rrp43L2
(Disordered)
Rrp45Cterm
State 1
Intermediate
State 2
Side View
Rrp6
S1/KH ring
PH-like ring
Central channel
Rrp44 PIN
Rrp44 CSD1
RNB channel-independent
RNB channel-dependent
L1
L2
S1
CSD2
RNB
90°
Top View
S1/KH ring
B
Rrp43L1
Rrp45Cterm
Rrp43L2
Rrp45
Rrp44
3'
5'
SO4 2-
Rrp43L2 termini
Rrp44 CSD1
C
Rrp43L1 termini

Figure 12. Order-disorder transitions in Rrp43 features and Rrp45$_{Cterm}$. (A through F) Structures were aligned based on the Rrp43 and Rrp45 chains or, in panel C, only Rrp45. RNase PH-like subunits Rrp41, Rrp46, Rrp42, and Mtr3 subunits are shown as a transparent gray surface for clarity. Terminal residues modeled for Rrp43L1 and Rrp43L2 features are shown as spheres and highlighted with arrows in the channel-dependent (D and E) and channel-independent (A, B and C) conformations, respectively. The dashed line in panel E indicates the presumed RNA path in that structure. **(**G) Rrp43L1 in the Exo11$^{44/6}$ structure. Mesh is a simulated annealing 2Fo – Fc map contoured at 1.5 σ and carved to within 3 Ångstroms of Rrp43. H) Rrp43L2 termini in the Exo11 structure. Mesh is a simulated annealing 2Fo – Fc map contoured at 1.0 σ. I) Panel H with Rrp43 from the Exo10^{44} structure (PDB 4IFD) aligned to the Rrp43 chain of Exo11$^{44/6}$ with electron densities shown as in panel H.


A
Exo11^44/6ΔNΔC
3'-3' 34 nt RNA (this work)
Rrp45
Rrp43
Rrp45Cterm
E304
Rrp43L1
Rrp43L2
termini
3'
Rrp44
(CSD1)
(PIN)
B
Exo12^44/6ΔC/47ΔC
18 nt AU RNA (PDB 5C0W)
E300
3'
C
Rrp44-Rrp41-Rrp45
No RNA (PDB 2WP8)
R303
D
Exo10^44/6Cterm
StemloopU30 RNA (PDB 4IFD)
Rrp43L1
termini
Rrp45Cterm
Rrp43L2
3'
E
Exo12^44/6ΔC/47ΔC
StemloopU30 RNA (PDB 5C0X)
3'
F
Exo10^6ΔC
PolyA24 RNA (PDB 4OO1)
G
Rrp43L1 (this work)
Rrp45
Rrp44
H
Rrp43L2 termini (this work)
Rrp43
I
(H) with Rrp43 from PDB 4IFD

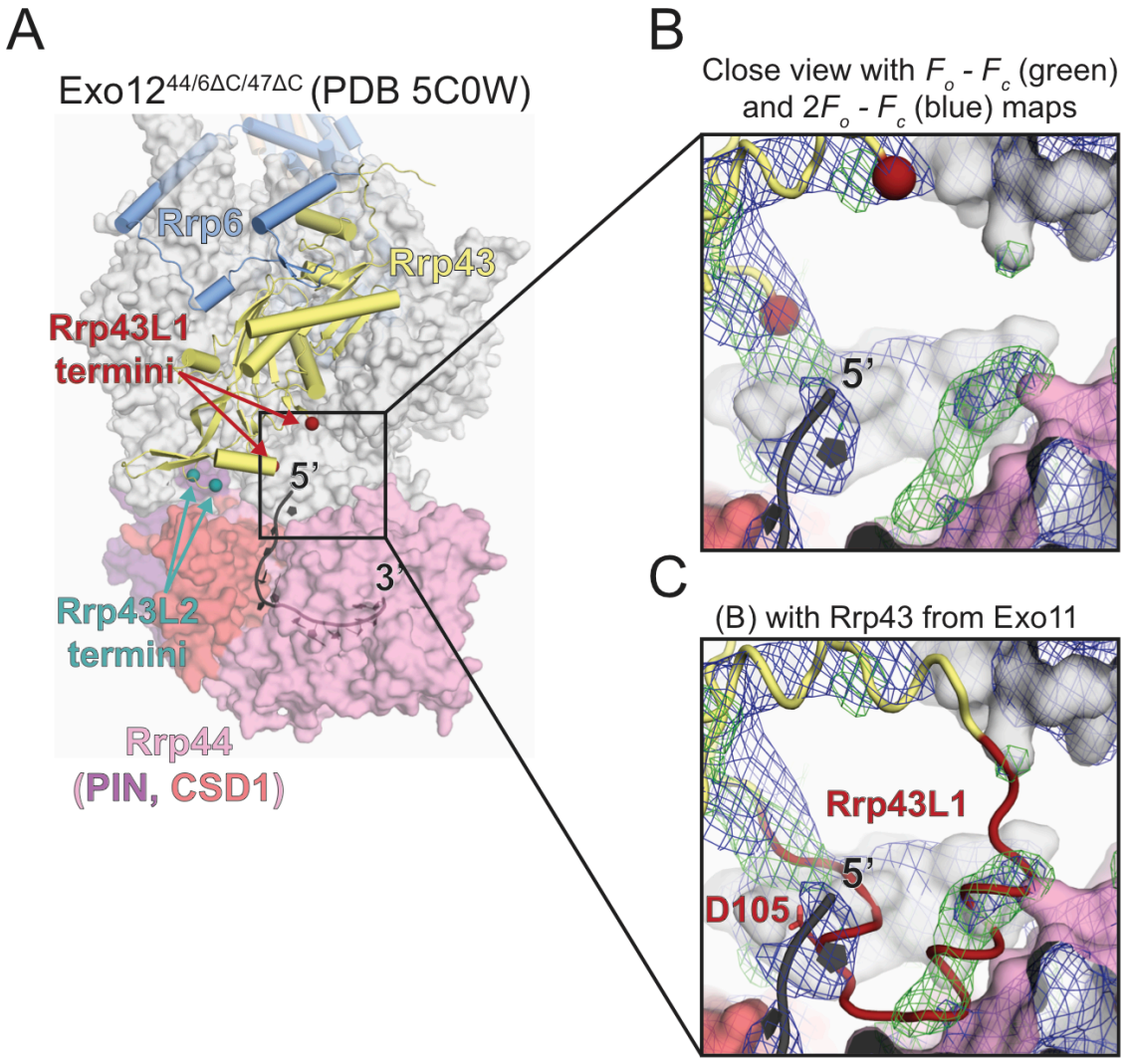


Figure 13. Rrp43L1 clashes with the 5' end of Rrp44 bound RNA from the Exo12$^{44/6/47}$ (A) Orientation of Exo12$^{44/6/47}$ used in analysis. (B) Close-up of RNA 5' end with Fo – Fc map shown as green mesh contoured at 2.0 σ and 2Fo – Fc map shown as blue mesh contoured to 1.0 σ. Rrp43 residues Thr99 and Asn126, the boundaries of Rrp43L1 in that model, are shown as spheres. Maps were generated in Phenix using the twin operator -h,-k, l. (C) The same view as in (B) but with Rrp43 replaced with Rrp43 from the Exo11$^{44/6}$ structure presented in this work. Alignment was performed based on the Rrp43 chain.

Comparison to the Exo10^{44}-RNA channel-dependent structure (PDB 4IFD) reveals that contacts between Rrp44 and the Rrp45$_{Cterm}$ and Rrp43L1 element are lost. While the Rrp43L1 element becomes disordered, the Rrp45$_{Cterm}$ forms a helix and does not contact Rrp44 in the channel-dependent Exo10^{44} (Liu et al., 2016; Makino et al., 2013b) or in the absence of Rrp44 in the Exo10^{6} structure (Wasmuth et al., 2014) (Figures 11A, 11C, and 12). Instead, different contacts are established between the Exo9 core and Rrp44 including Rrp43 residues 251-270 (Rrp43L2), which form a beta hairpin that bridges the Rrp44 CSD1, catalytic domain, and RNA (Liu et al., 2016; Makino et al., 2013b) (Figures 11A, 11C, and 12). By contrast, this feature of Rrp43 is disordered in our structure (Figures 11A, 11B, and 12). A recent structure of Exo11$^{44/Ski7N\text{-}term}$ with Rrp44 bound in the through-channel conformation (PDB 5JEA) observed that Rrp43L2 was disordered (Kowalinski et al., 2016), but the authors noted that Rrp44 CSD1 is rotated relative to other RNA-through-channel structures due to crystal lattice contacts in that region, suggesting that ordering of Rrp43L2 may be dependent on rotation of Rrp44 CSD1 to that observed in the channel-dependent Rrp44 conformation. The unique contacts between Rrp44 and the Exo9 core in channel-dependent and channel-independent Rrp44 conformations suggested they might contribute differentially to the ability of Rrp44 to bind and degrade RNA.

We reconstituted exosome complexes containing deletions of Rrp43L1, Rrp43L2, and/or deletion of Rrp45$_{Cterm}$ (Figures 14, 15, 16A, and 16) to test the importance of these elements in RNA binding and degradation assays. Mutant Rrp43 and Rrp45 behaved as wild-type with respect to formation of Exo9 complexes and stable association with Rrp44. Consistent with Rrp43L1 and Rrp43L2 contacts to Rrp44 and not Rrp6, no differences in activity were observed for Rrp6 when associated with Exo9 containing the Rrp43 mutants (Figures 16C, and 16D).

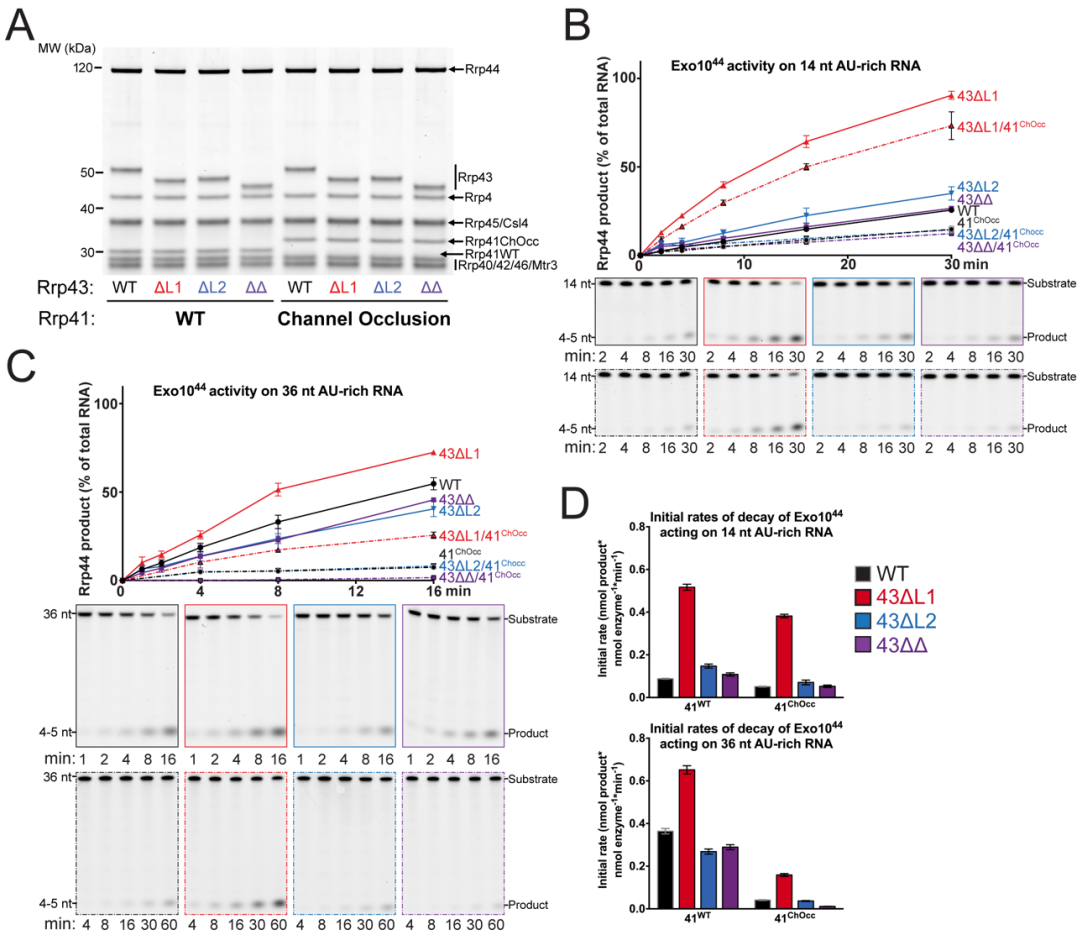


Figure 14. Exo10^{44} activity on AU-rich RNAs is enhanced by deletion of Rrp43L1. (A) SDS-PAGE analysis of reconstituted Exo10^{44} complexes containing Rrp43 mutants. '41^{ChOcc}' refers to a loop insertion mutant in Rrp41 that occludes the central channel. (B and C) WT and mutant Exo10^{44} activity on 14 nt (B) and 36 nt (C) 5' FAM AU-rich RNA substrates. The experiment was performed in triplicate and quantification of mean values is shown with error bars at ±1 standard deviation. Representative polyacrylamide TBE-urea gels are shown below. (D) Quantification of initial rates of decay of Exo10^{44} complexes on 14 nt and 36 nt 5' FAM AU-rich RNA substrates from the experiments in panels (B) and (C). Mean values of the triplicate experiments are plotted with error bars at ±1 standard deviation.

Figure 15. Deletion of features associated with the channel-independent conformation of Rrp44 enhance its activity in Exo11$^{44/6}$ complexes. (A) SDS-PAGE analysis of reconstituted Exo11$^{44/6}$ complexes containing mutant Rrp43 and/or Rrp45 proteins. (B and C) WT and mutant Exo11$^{44/6}$ activity on 14 nt (B) and 37 nt (C) 5' FAM poly(A) RNA substrates. The experiment was performed in triplicate and quantification of mean values is shown with error bars at ±1 standard deviation. Representative polyacrylamide TBE-urea gels are shown below. (D) Initial rates of decay of Rrp44 in Exo11$^{44/6}$ complexes and free Rrp44 on short (14 nt) and long (37 nt) 5' FAM poly(A) RNA substrates from the experiments in panels (B) and (C). Mean values are shown with error bars at ±1 standard deviation. (E and F) WT and mutant Exo11$^{44/6}$ activity on 14 nt (E) and 36 nt (F) AU-rich RNA substrates. The experiment was performed in triplicate and quantification of mean values is shown with error bars at ±1 standard deviation. Representative polyacrylamide TBE-urea gels are shown below.

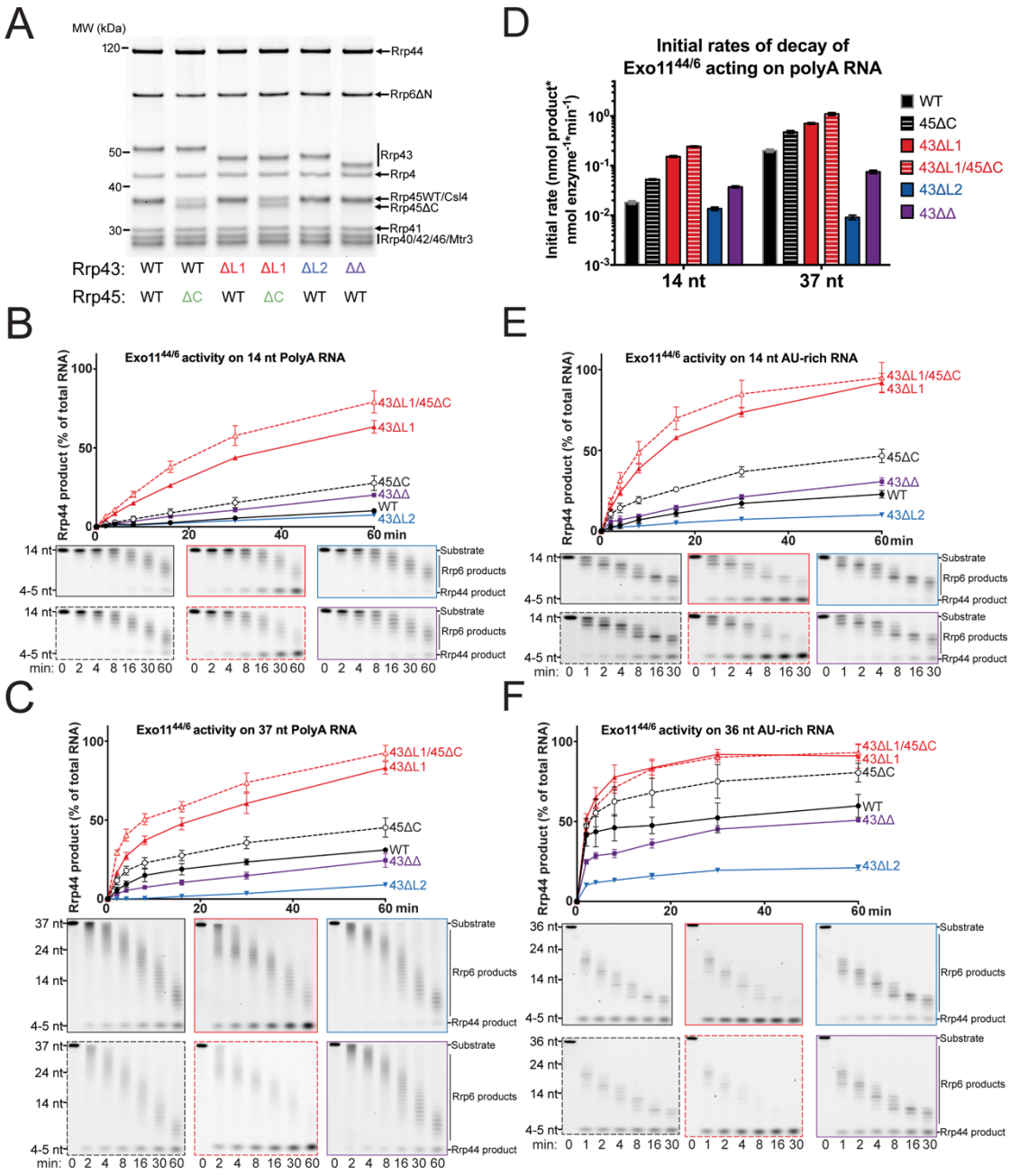
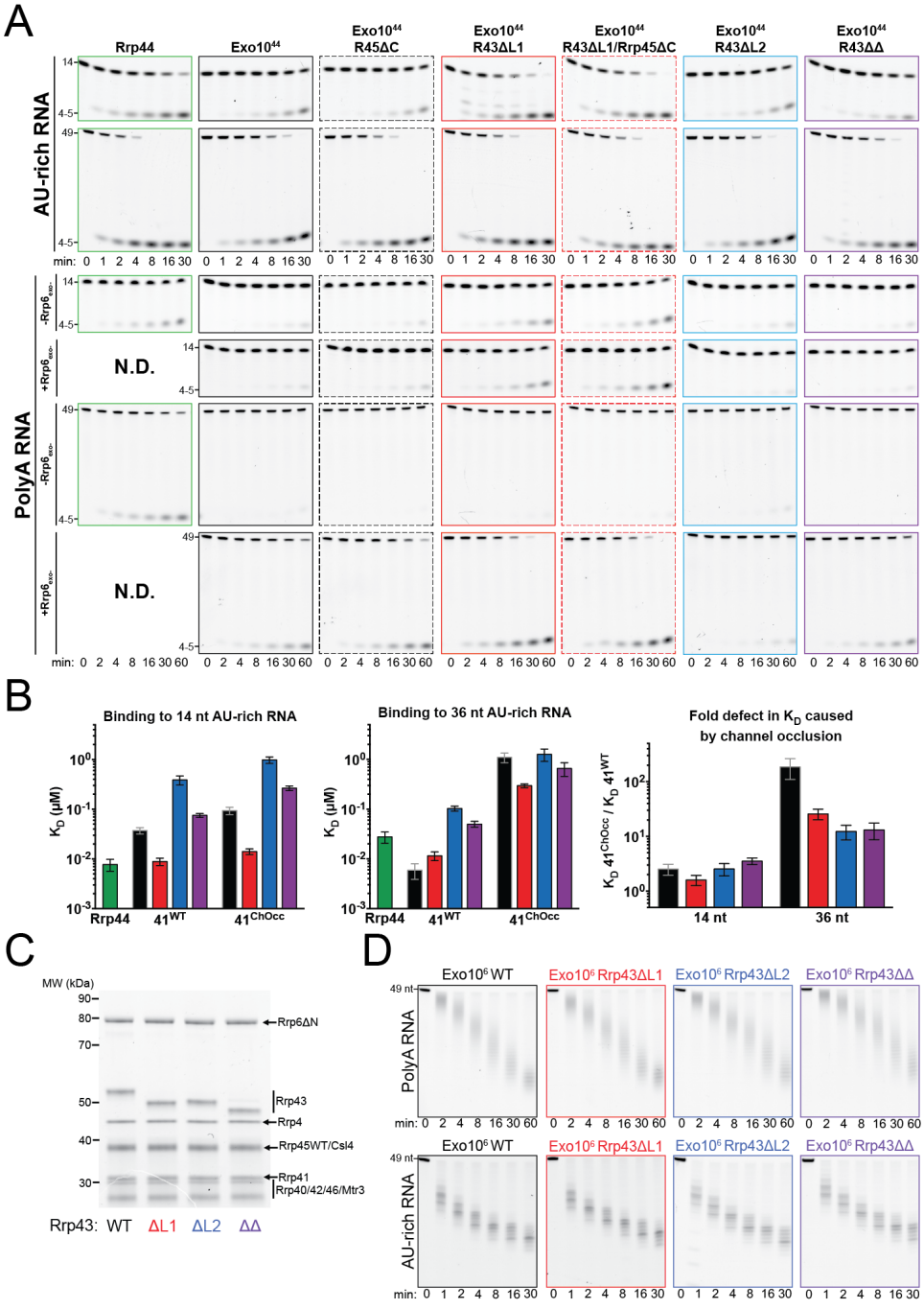

A
MW (kDa)
120
50
40
30
Rrp44
Rrp6ΔN
Rrp43
Rrp4
Rrp45WT/Csl4
Rrp45ΔC
Rrp41
Rrp40/42/46/Mtr3
Rrp43: WT WT ΔL1 ΔL1 ΔL2 ΔΔ
Rrp45: WT ΔC WT ΔC WT WT
D
Initial rates of decay of
Exo11^44/6 acting on polyA RNA
Initial rate (nmol product* nmol enzyme^-1*min^-1)
WT
45ΔC
43ΔL1
43ΔL1/45ΔC
43ΔL2
43ΔΔ
14 nt
37 nt
B
Exo11^44/6 activity on 14 nt PolyA RNA
Rrp44 product (% of total RNA)
43ΔL1/45ΔC
43ΔL1
45ΔC
43ΔΔ
WT
43ΔL2
60 min
14 nt
4-5 nt
Substrate
Rrp6 products
Rrp44 product
min: 0 2 4 8 16 30 60
E
Exo11^44/6 activity on 14 nt AU-rich RNA
Rrp44 product (% of total RNA)
43ΔL1/45ΔC
43ΔL1
45ΔC
43ΔΔ
WT
43ΔL2
60 min
14 nt
4-5 nt
Substrate
Rrp6 products
Rrp44 product
min: 0 1 2 4 8 16 30
C
Exo11^44/6 activity on 37 nt PolyA RNA
Rrp44 product (% of total RNA)
43ΔL1/45ΔC
43ΔL1
45ΔC
WT
43ΔΔ
43ΔL2
60 min
37 nt
24 nt
14 nt
4-5 nt
Substrate
Rrp6 products
Rrp44 product
min: 0 2 4 8 16 30 60
F
Exo11^44/6 activity on 36 nt AU-rich RNA
Rrp44 product (% of total RNA)
43ΔL1/45ΔC
43ΔL1
45ΔC
WT
43ΔΔ
43ΔL2
60 min
36 nt
24 nt
14 nt
4-5 nt
Substrate
Rrp6 products
Rrp44 product
min: 0 1 2 4 8 16 30

Figure 16. Activities of Exo10^{6} and Exo10^{44} complexes containing wild-type and mutant Rrp43 and Rrp45 proteins. (A) Exo10^{44} decay activity on 5' FAM 14 nt and 49 nt AU-rich and poly(A) substrates. Where indicated, a 3-fold molar excess (3 nM final concentration) of a catalytically inactive Rrp6 protein (residues 129-733, D238N mutant) was added to the reaction. The full-length RNA substrate is indicated with 14 nt or 49 nt marker and the Rrp44 decay product is indicated with the 4-5 nt marker. N.D. indicates panels that were not done. B) Binding affinities of Exo10^{44} complexes and free Rrp44 for AU-rich RNA as measured by fluorescence anisotropy. '41^{ChOcc}' refers to a mutant in Rrp41 that occludes the central channel. Error bars indicate ±1 standard deviation obtained from three independent experiments. C) SDS-PAGE analysis of reconstituted Exo10^{6} complexes containing deletions in Rrp43. Gel is 4-12% acrylamide BIS-TRIS run in MOPS-SDS running buffer and stained with SYPRO ruby. D) Decay time courses of reconstituted Exo10^{6} complexes on 5' FAM poly(A) and AU-rich 49 nt RNA.


A
Rrp44
Exo10^44
Exo10^44 R45ΔC
Exo10^44 R43ΔL1
Exo10^44 R43ΔL1/Rrp45ΔC
Exo10^44 R43ΔL2
Exo10^44 R43ΔΔ
AU-rich RNA
PolyA RNA
-Rrp6exo-
+Rrp6exo-
N.D.
min: 0 1 2 4 8 16 30
min: 0 2 4 8 16 30 60
B
Binding to 14 nt AU-rich RNA
Binding to 36 nt AU-rich RNA
Fold defect in K_D caused by channel occlusion
K_D (μM)
Rrp44
41^WT
41^ChOcc
14 nt
36 nt
C
MW (kDa)
Rrp6ΔN
Rrp43
Rrp4
Rrp45WT/Csl4
Rrp41
Rrp40/42/46/Mtr3
Rrp43: WT ΔL1 ΔL2 ΔΔ
D
Exo10^6 WT
Exo10^6 Rrp43ΔL1
Exo10^6 Rrp43ΔL2
Exo10^6 Rrp43ΔΔ
PolyA RNA
AU-rich RNA
49 nt

We first assessed the impact of deleting Rrp43L1 and Rrp43L2 in the context of $Exo10^{44}$ complexes in degradation of long (36/37 nt) and short (14 nt) AU-rich and poly(A) RNA substrates (Figures 14 and 16A). Deleting Rrp43L1 increased Rrp44 activity on short and long AU-rich RNA substrates. In contrast, deletion of Rrp43L2 had no detectable impact on degradation of short AU-rich RNA although a measurable defect was observed in degrading the long AU-rich substrate (Figures 14B and 14C). Interestingly, deletion of Rrp43L2 suppressed the stimulation observed for Rrp43L1 deletion in degradation assays with short and long RNA when Rrp43L1 and Rrp43L2 deletions were combined. A similar pattern was observed for degradation of short poly(A) RNA for the respective $Exo10^{44}$ complexes, however none of these complexes exhibited activity on 37 nt poly(A) RNA in comparison to Rrp44 (Figure 16A). The latter observation is consistent with previous results showing that the Rrp6 protein was required for Rrp44 to bind and degrade poly(A) RNA in nuclear exosome complexes (Wasmuth and Lima, 2012a).

Binding and decay assays with the AU-rich RNA substrates were also performed using $Exo10^{44}$ complexes that contained a loop-insertion mutant of Rrp41 (41^{ChOcc}) that partially occludes the central channel (Wasmuth and Lima, 2012a). We observed that binding and initial rate of decay were inhibited when the channel was occluded, supporting the conclusion that Rrp44 remains dependent on RNA passing through the central channel when bound to Exo9 cores containing the mutated Rrp43 subunit (Figure 16B). Notably, the dependence on the central channel was more pronounced for the 36 nt AU-rich RNA than the 14 nt RNA (Figure 16B), consistent with previous observations that Rrp44 conformational changes from channel-independent to the channel-dependent are promoted by incubation with long (>24 nt) single-stranded RNAs under non-degrading conditions (Liu et al., 2014).

We next assessed reconstituted wild-type and mutant nuclear Exo11$^{44/6\Delta N}$ complexes in degradation assays using short (14 nt) or long (37/36 nt) poly(A) and AU-rich RNAs (Figure 15). For Exo11$^{44/6\Delta N}$ complexes, deletion of Rrp43L1 increased Rrp44 activity for each RNA substrate tested, especially evident for the short RNA substrates. In contrast to that observed for Exo10^{44} complexes, deletion of Rrp43L2 resulted in a measurable loss of Rrp44 activity for the 14 nt AU-rich while deletion of Rrp43L2 had little impact on degradation of the 14 nt poly(A) RNA. Deletion of Rrp43L2 reduced Rrp44 activity on both 36 nt AU-rich and 37 nt poly(A) substrates. Consistent with results obtained using Exo10^{44} complexes, deletion of Rrp43L2 suppressed the stimulatory effect of removing Rrp43L1 while removal of Rrp43L1 largely suppressed the defects observed for Rrp43L2 deletion. Deletion of Rrp45$_{Cterm}$ tail resulted in a measurable increase in Rrp44 activity, albeit less than that observed for deletion of Rrp43L1. Deletion of both the Rrp45$_{Cterm}$ tail and Rrp43L1 appeared additive. The presence of Rrp6 activities in these assays could lead to artifacts in analysis of Rrp44 decay products, so we also added catalytically inactive Rrp6 (Rrp6$_{exo-}$) to Exo10^{44} complexes in degradation assays with the poly(A) RNA substrates (Figure 16A). Results from these assays are qualitatively similar to those obtained for catalytically active Exo11$^{44/6\Delta N}$ complexes. Deletion of Rrp43L1 and Rrp45$_{Cterm}$ stimulated degradation, deletion of Rrp43L2 inhibited degradation, and deleting both Rrp43L1 and Rrp43L2 suppressed the effects observed for the individual deletions.

Taken together, these results suggest that Exo9 subunit contacts between the Rrp43L1, Rrp45$_{Cterm}$ and channel-independent Rrp44 conformation inhibit Rrp44 activity in Exo10^{44} and Exo11$^{44/66\Delta N}$, while Rrp43L2 contacts to the channel-dependent Rrp44 conformation promote activity in Exo11$^{44/6}$. Longer RNA substrates presumably enter the 11-component exosome via its central channel, and the greater

dependence on Rrp43L2 for degrading longer RNA is consistent with Rrp43L2 contacting RNA as it exits the central channel to Rrp44 (Liu et al., 2016; Makino et al., 2013b). Rrp43L2 appears dispensable for degradation of shorter RNA substrates, but enhanced degradation of 14 nt RNA by deletion of Rrp43L1 and $Rrp45_{Cterm}$ was suppressed by deletion of Rrp43L2. Although other models are possible (see Discussion), we believe the simplest explanation is that Rrp44 is stabilized in the channel-independent conformation by contacts to Rrp43L1 and $Rrp45_{Cterm}$, and that Rrp44 is stabilized in the channel-dependent conformation by contacts to Rrp43L2. Because deletion of Rrp43L1 and $Rrp45_{Cterm}$ enhances activity in a manner dependent on Rrp43L2, we posit that Rrp44 conformational changes are important for degrading both long and short substrates independent of whether RNA enters the exosome through the central channel (long RNA) or via direct binding to Rrp44 (short RNA). It remains unclear if this could be a regulated feature of the RNA exosome, or if it serves to inhibit Rrp44 activities when RNA is not presented through the Exo9 central channel, the presumed route for most cellular RNAs (Delan-Forino et al., 2017; Drazkowska et al., 2013; Wasmuth and Lima, 2012a).

In vitro analysis of a Dis3 direct-access mutant

In a recent study, investigators analyzed the in vivo effects in a strain expressing a *rrp44* allele that harbored five alanine substitutions predicted to destabilize the channel-independent Rrp44 conformation but not affect the channel-dependent conformation (Han and van Hoof, 2016) (Figure 17). Among many observations, they showed that cells expressing this *rrp44* allele were viable, albeit with a slow-growth phenotype (Figure 18A) and defects in degradation of structured substrates. The mutant, which they termed Rrp44da, also showed synthetic defects when combined with mutations that rendered the endo- and exoribonuclease sites of

Rrp44 inactive (Figure 18A). Perhaps most intriguing was the observation the *rrp44da* allele suppressed growth defects/lethality of mutations designed to occlude the Exo9 central channel (Wasmuth and Lima, 2012a), suggesting strong genetic interactions between mutations in the Exo9 core and Rrp44 (Figures 18B and 18C).

Because their study lacked in vitro characterization of the mutant, we cloned (original clone obtained from A. van Hoof), expressed, purified, and reconstituted Rrp44da into nuclear exosome complexes, and assayed its decay activities (Figure 19). Rrp44da expressed, purified, and reconstituted into an Exo12$^{44/6/47}$ complex as WT, indicating that the mutations did not result in severe protein folding defects (Figures 19A and 19B). We assayed the mutant enzyme in and out of the complex for decay activity compared to the WT (Figure 19C). For both the free enzyme and the complex we observed markedly decreased intensity of Rrp44 products on 14 and 49 nt substrates relative to WT for both poly(A) and AU-rich RNAs. Because the mutations were designed to specifically inhibit conformational changes in the context of the exosome complex, the decrease in activity of the free enzyme was unexpected. However, the synthetic lethality of combining Rrp44exo- inactivation with these mutations and the rescuing of the channel block phenotype suggests that this mutant is not simply decreasing Rrp44 activity in vivo (Figure 18). That being said, mutation of D602, which forms a salt bridge to R600 near the 3' end of the bound RNA, could be responsible for diminishing the activity of this enzyme (see next section). Additionally, the mutation of two consecutive arginines to alanines results in a considerable loss of steric bulk at that location, which could affect local folding or conformational dynamics important for substrate binding.

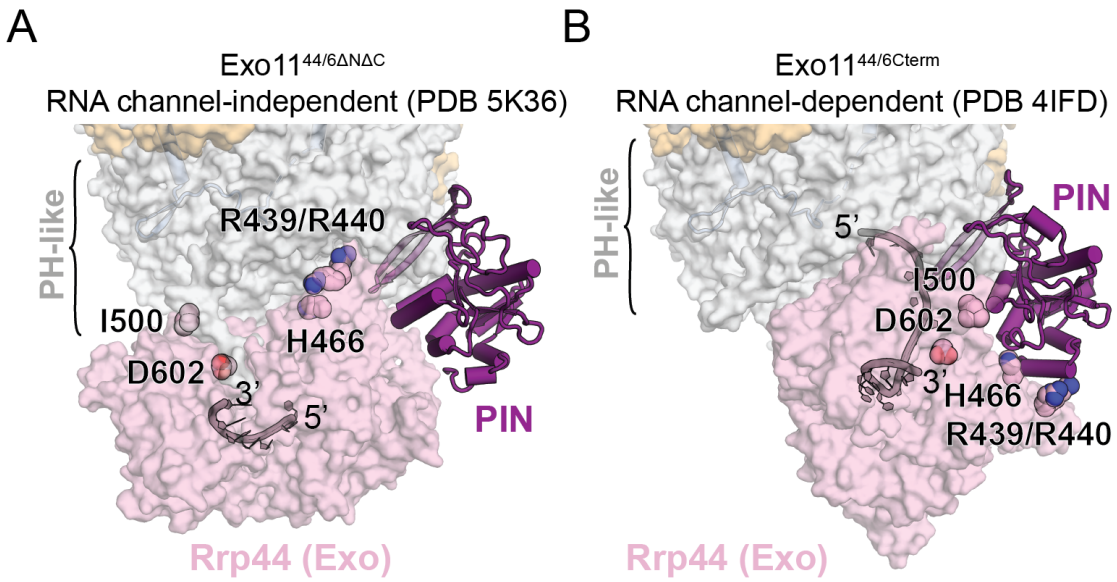


Figure 17. Structural modeling of Rrp44da mutations. Side-chains of the residues mutated to alanine in the Rrp44da mutant are highlighted as spheres in the RNA channel-independent (A) and RNA channel-dependent (B) conformations of *S. cerevisiae.* Structures were aligned based on all six PH-like proteins and RNA 3' and 5' ends are indicated. 'Rrp44 (Exo)' refers to the exonuclease module of Rrp44, which encompasses its two cold shock domains, its RNB catalytic domain and its S1 domain.

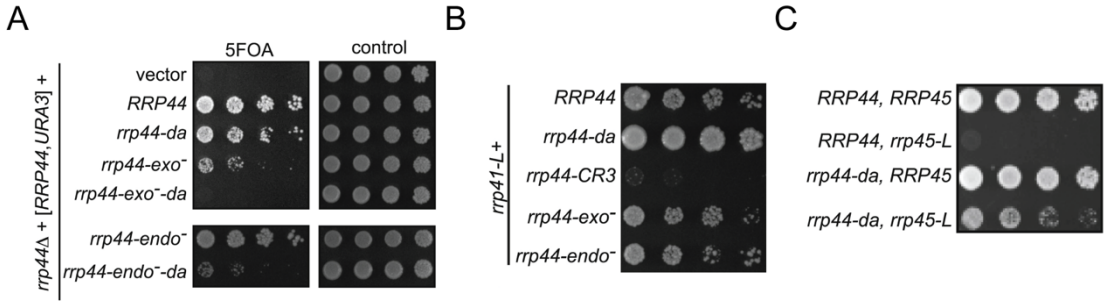


Figure 18 The Rrp44da-exosome uses both the exo- and endonuclease activities and suppresses growth defects of channel occlusions. (A) *rrp44-da* is synthetic lethal with *rrp44-exo*$^{-}$ and *rrp44-endo*$^{-}$. An *rrp44*Δ strain carrying a wild-type *RRP44* allele in a *URA3* plasmid was transformed with *LEU2* plasmids carrying *RRP44* variants. The transformants were serially diluted and spotted on 5FOA and SC-LEU-URA (control) media. (B and C) *rrp44-da* suppresses the growth defect of *rrp41-L* and *rrp45-L*. *rrp44*Δ *rrp41*Δ or *rrp44*Δ *rrp45*Δ strains carrying *RRP44* variants in a *LEU2* plasmid and a wild-type *RRP41* or *RRP45* in a *URA3* plasmid were transformed with a *TRP1* plasmid carrying the *rrp41-L* or *rrp45-L* allele. The transformants were serially diluted and spotted on 5FOA media. Figure and caption adapted from Figures 3 and 5 of Han and van Hoof, 2016.

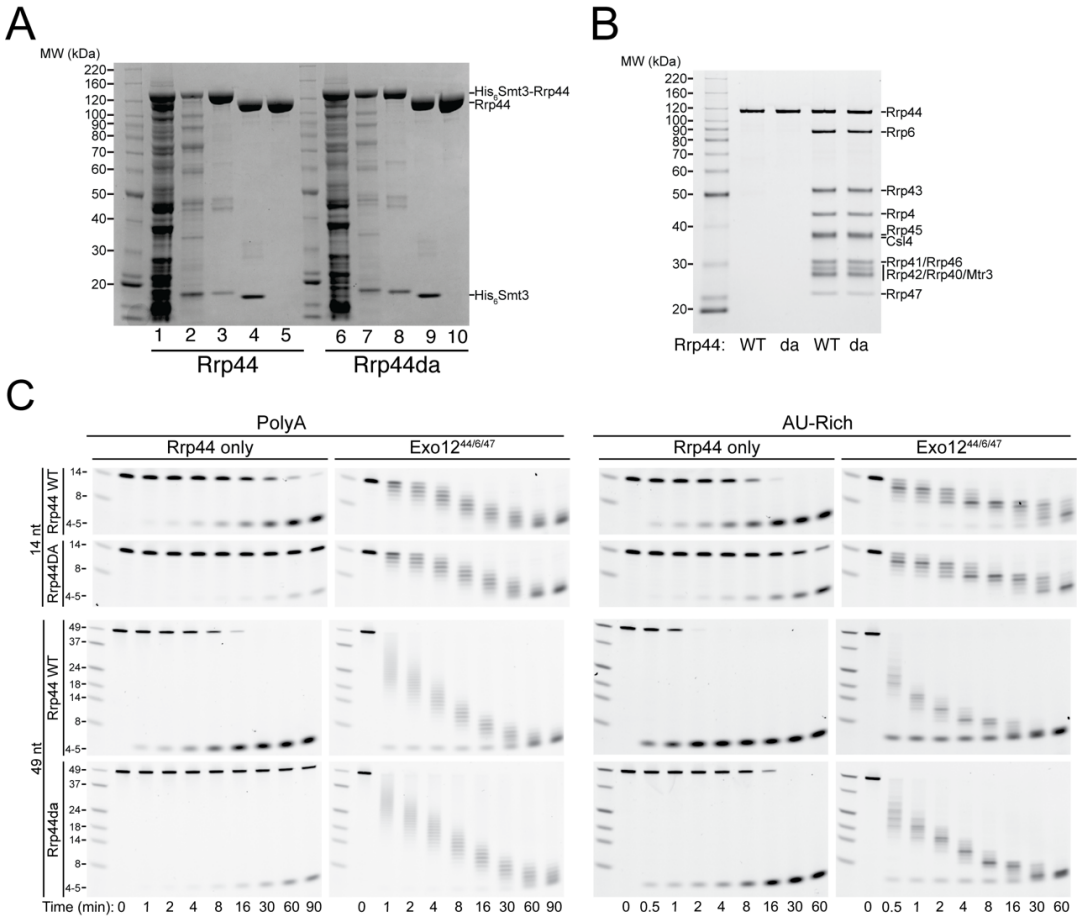
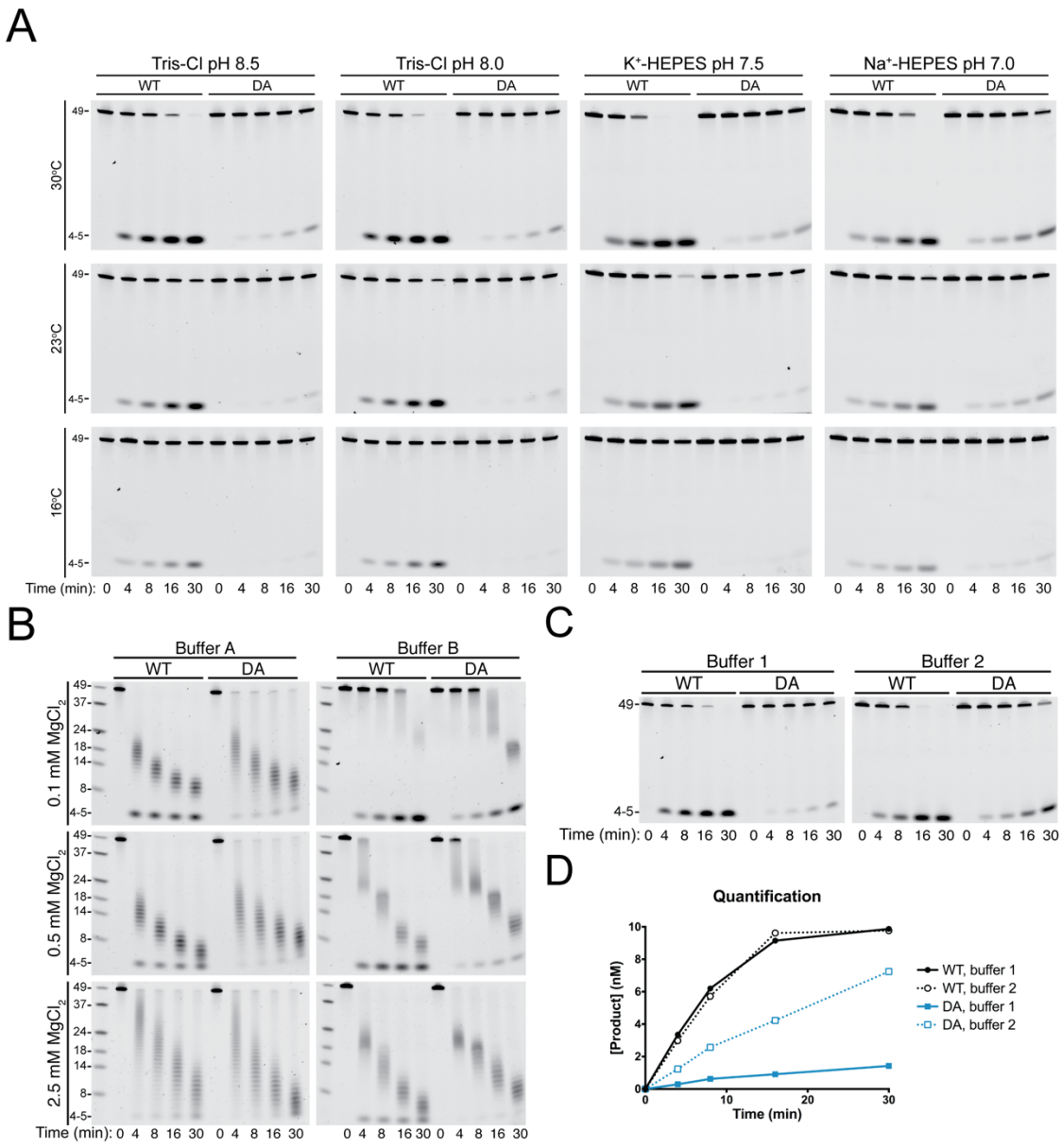


Figure 19: Rrp44da mutant has decreased activity on a variety of RNA substrates under standard assay conditions. (A) SDS-PAGE analysis of parallel preparations of WT and da mutant Rrp44 enzymes. Lanes 1 and 6 are whole cell extracts 16 hrs after induction with IPTG, lanes 2 and 7 are eluates from the Ni-NTA column, lanes 3 and 8 are eluates from the heparin column, lanes 4 and 9 are heparin eluates post Ulp treatment, and lanes 5 and 10 are eluates from gel filtration. Gel was stained with Coomassie brilliant blue for visualization. (B) Protein gel showing purified free enzymes (first two lanes after marker) and gel-filtration purified Exo12$^{Rrp6/Rrp47/Rrp44}$ complexes (3rd and 4th lanes after marker). Gels were stained with SYPRO Ruby for visualization. (C) RNA decay activities of Rrp44 free enzymes and nuclear exosome complexes. Substrates are as indicated and contain a 5' fluorescein label. Gels are 15% polyacrylamide TBE-urea and were imaged for fluorescein signal.

In spite of these issues, buffer pH, [Mg^{2+}], and reaction temperature were screened in an effort to find conditions that mitigated the activity discrepancy (Figure 20). From these efforts, we found that a reaction buffer that contains a lower [Mg^{2+}] and is at a lower pH than our standard buffer retains activity of the wild-type enzyme but has substantially higher activity for the da mutant (Figures 20C and 20D).

Rrp44 da mutant and WT were reconstituted into complexes containing a loop insertion in Rrp41 (41^{ChOcc}, see above) and were assayed for degradation activity on poly(A) RNA in the original buffer (Figure 21). We determined that Rrp44 activity was highly compromised in the Rrp44da/$Rrp41^{ChOcc}$ complex under these conditions (Figure 21). Though this experiment bears repeating using the newly formulated buffer conditions, these results are still at odds with the genetic findings (Figure 19B), and suggest that a different approach will be required in future studies to de-convolute the in vivo effects of Rrp44 activity loss and perturbation of conformational state (see Discussion).

Figure 20: Rrp44da favors low pH and low magnesium. All gels are 15% acrylamide TBE-urea and were imaged for fluorescein signal. (A) Decay of 5' fluorescein poly(A) 49 nt RNA by Rrp44 using different buffer pH and incubation temperature. The reaction solution contained 50 mM KCl, 0.5 mM $MgCl_2$, 1 mM TCEP HCl, 0.05 % v/v IGEPAL and the indicated buffer was included at 20 mM. Final pHs of the solutions were also measured empirically and were as follows: 8.3, 7.8, 7.1, 5.9. (B) Decay of poly(A) 49 nt RNA by Rrp44 under varying $MgCl_2$. Base buffer was 20 mM HEPES-KOH pH 7.5, 50 mM KCl, 1 mM TCEP pH 7.0, 0.05 % v/v IGEPAL and was at a final pH of 7.3. (C) Decay of 5' fluorescein poly(A) 49 nt RNA by Rrp44. Buffer 1 contained 20 mM Tris-Cl pH 8.0, 50 mM KCl, 0.5 mM $MgCl_2$, 1 mM TCEP-HCl, 0.05 % v/v IGEPAL and was at a final pH of 7.8. Buffer 2 contained 20 mM HEPES-NaOH pH 7.0, 50 mM NaCl, 0.1 mM $MgCl_2$, 1 mM TCEP pH 7.0, 0.05 % v/v IGEPAL and was at a final pH of 6.8. (D) Superficial quantification of product for the experiment in panel B. Conversion to nM was achieved by subtracting the fraction of substrate remaining at the given incubation time from 1 and multiplying that number by 10. Because only a single replicate was performed, no error bars are shown.


A
Tris-Cl pH 8.5
Tris-Cl pH 8.0
K+-HEPES pH 7.5
Na+-HEPES pH 7.0
WT
DA
30°C
23°C
16°C
49-
4-5-
Time (min): 0 4 8 16 30
B
Buffer A
Buffer B
0.1 mM MgCl2
0.5 mM MgCl2
2.5 mM MgCl2
49-
37-
24-
18-
14-
8-
4-5-
C
Buffer 1
Buffer 2
D
Quantification
[Product] (nM)
Time (min)
WT, buffer 1
WT, buffer 2
DA, buffer 1
DA, buffer 2

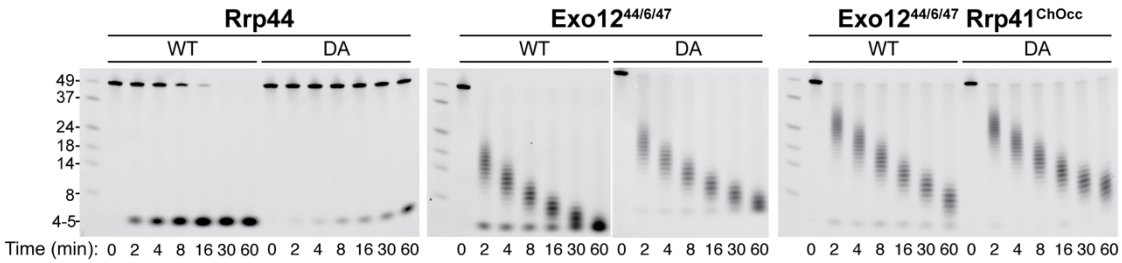


Figure 21: Rrp44da fails to rescue activity of a channel occluded nuclear exosome in vitro. Decay of 5' fluorescein poly(A) 49 nt RNA by Rrp44 and nuclear exosome complexes using the original RNA decay buffer (20 mM Tris-Cl pH 8.0, 50 mM KCl, 0.5 mM $MgCl_2$, 1 mM TCEP-HCl, 0.05 % v/v IGEPAL). '41ChOcc' indicates that the complex contains a loop insertion in Rrp41 designed to occlude the central channel. Gels are 15% acrylamide TBE-urea and were imaged for fluorescein signal.

3' phosphorylated RNA is a Rrp44 substrate

The RNA path for the last 7 nt of RNA in the Rrp44 active site in our structure resembles that previously observed in the structure of Rrp44 bound to poly(A) RNA (Lorentzen et al., 2008). A similar RNA path was modeled in the nuclear RNA exosome structure bound to an 18 nt AU-rich RNA (Makino et al., 2015), but its path diverges after the 7th nt and its 5' end is modeled into densities that we assign to the Rrp43L1 loop near residue Asp105 (Figure 13). Inspection of electron densities proximal to the Rrp44 active site and 3' end of RNA in our structure revealed a sulfate ion that was present in the crystallization medium (Figure 22A). The sulfate ion is positioned 2.6 Å or 3.4 Å from the 2' and 3' OH groups of the terminal base, respectively, and is situated in a positively charged pocket formed by Rrp44 Arg600 (2.8 Å) and Rrp45 Arg303 (2.7 Å).

The location of the sulfate ion in the Rrp44-RNA complex suggested that Rrp44 could accommodate a 3' phosphorylated RNA in its active site. In contrast, inspection of the Rrp6 active site in this and previous structures of Rrp6 bound to AMP (PDB 2HBL) or Exo10$^{6exo\text{-}\Delta N\Delta C}$ bound to poly(A) (PDB 4OO1) suggested that the Rrp6 active site could not accommodate a 3' phosphorylated RNA because the RNA 2' OH is 2.5 Å from the backbone carbonyl of His 241 while the 3' OH is coordinated by side chains of Asp240 and backbone amide of His241 (Midtgaard et al., 2006; Wasmuth et al., 2014) (Figures 22B and 23). The RNA and last nucleotide adopts a different configuration in the twinned structure of Rrp6 catalytic domain bound to polyU RNA modeled with partial occupancy (PDB 5C0Y) (Figure 23), however contacts to the RNA 3' end in this structure would still appear to preclude accommodation of a 3' phosphorylated RNA substrate (Makino et al., 2015).

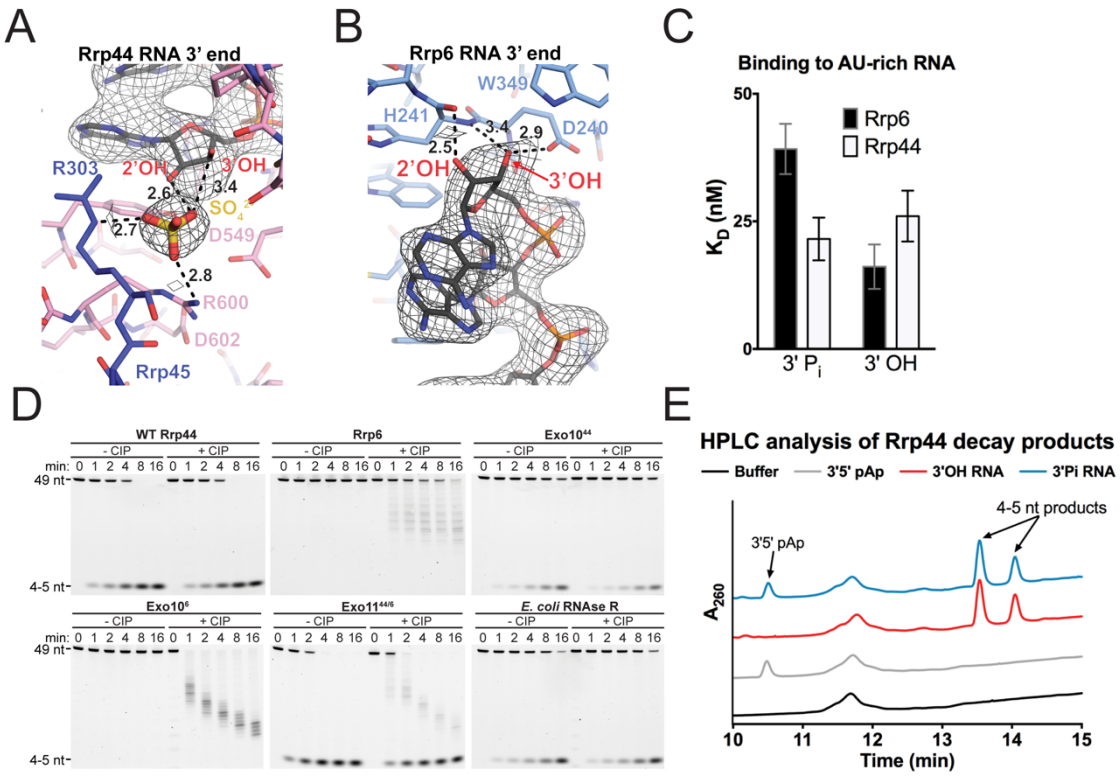


Figure 22. 3' Phosphate RNA is a substrate of Rrp44 but not Rrp6. (A and B) 3' end of the RNA bound to Rrp44 (A) and Rrp6 (B). Mesh represents electron density from a simulated annealing 2Fo-Fc omit map contoured at 1.2 σ. Distances are shown in Ångstroms. (C) Binding affinities of Rrp44 and Rrp6 for a 5' FAM 3' phosphate or 3' OH 36 nt AU-rich RNA as measured by fluorescence anisotropy. Mean values are plotted with error bars at ±1 standard deviation of a triplicate experiment. (D) Polyacrylamide TBE-urea gel analysis of RNA decay time-courses using a 5' FAM 3' phosphate AU-rich 49 nt RNA performed in the presence or absence of calf intestinal phosphatase (CIP) as indicated. (E) Ion-pair reverse-phase HPLC analysis of Rrp44 decay products. Reactions contained Rrp44 and either 5' FAM 3' phosphate 49 nt AU-rich RNA, 5' FAM 3' OH 49 nt AU-rich RNA, 3'5' pAp, or no substrate. Absorbance was monitored at 260 nm for detection of nucleotides and 4-5 nt RNA decay products.

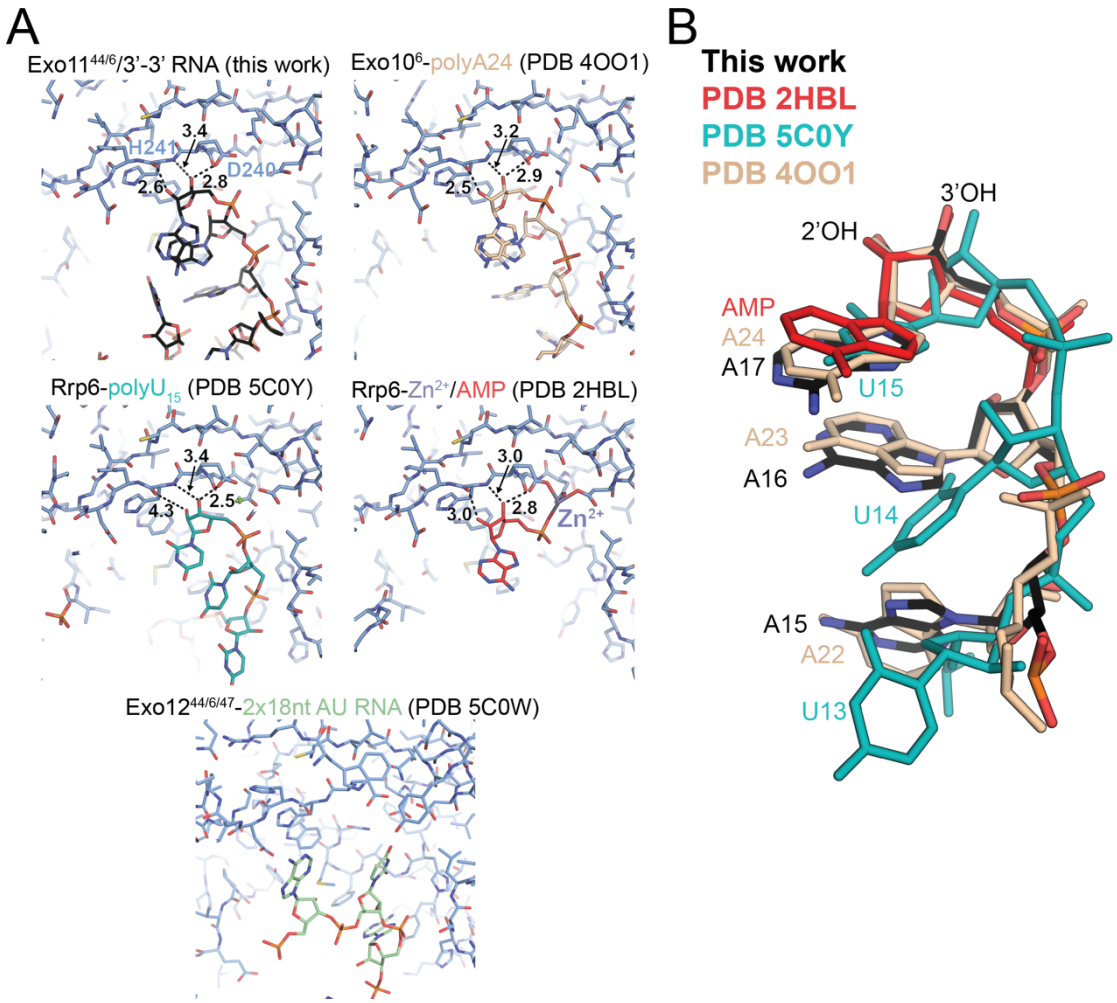
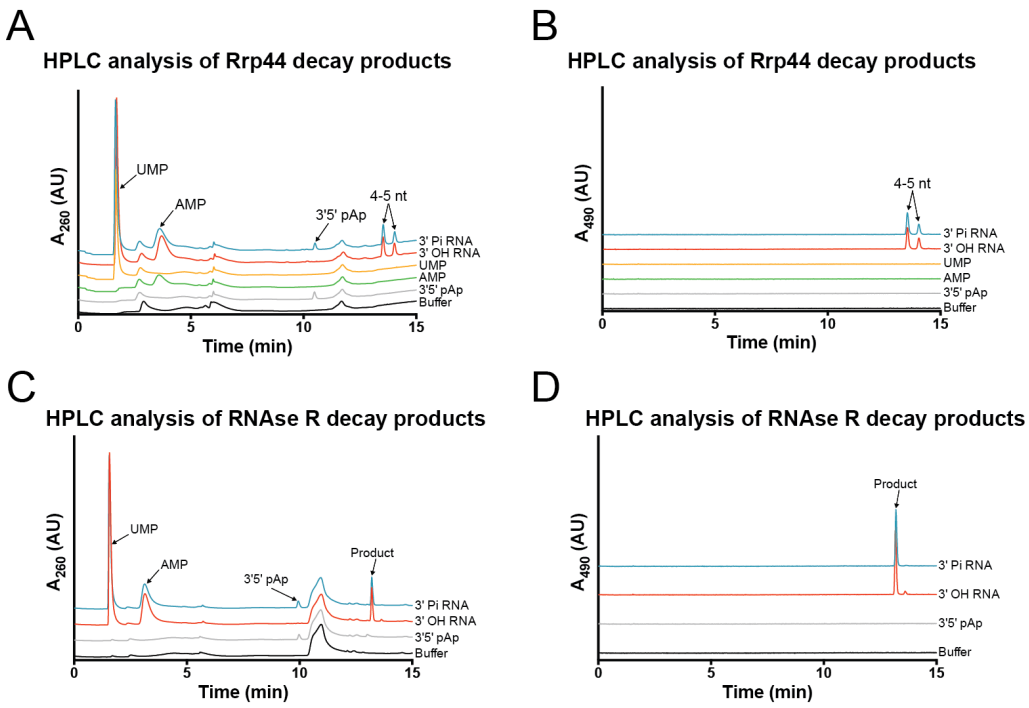


Figure 23. RNA in proximity to the Rrp6 active site. (A) RNA in or near the active site of Rrp6 in several structures in the PDB. Distances to the backbone carbonyl and amide of His241 and side chain of Asp240 are shown in Ångstroms. (B) The three terminal nucleotides of RNA or AMP in the Rrp6 active site from four different structures. Rrp6 residues 150-300, which span the catalytic residues, were aligned and the protein is removed for clarity.

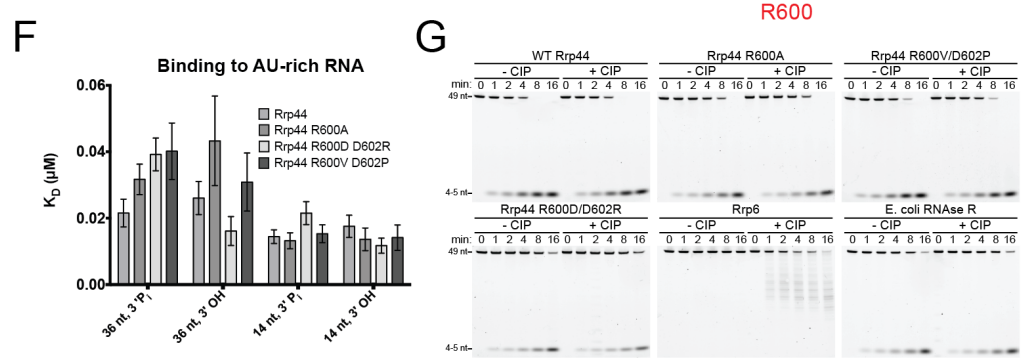
Consistent with the aforementioned predictions, Rrp44 can bind and degrade a 3' phosphorylated RNA (Figures 22C, 22D, 24F and 24G). By contrast, Rrp6 cannot degrade 3' phosphate RNA, consistent with previous findings (Burkard and Butler, 2000), despite its ability to bind it (Figures 22C, 22D, and 24G). These respective activities persist when Rrp6 and Rrp44 are reconstituted in complexes as $Exo10^{44}$, $Exo10^{6\Delta N}$ or $Exo11^{44/6\Delta N}$ (Figure 22D). A question remained as to whether Rrp44 initiates degradation of the 3' phosphate RNA substrate by first removing the 3' phosphate of the terminal ribonucleotide, followed by subsequent hydrolysis and release of NMP products, or if it accommodates the 3' phosphorylated terminal ribonucleotide in the active site, releasing 3'-5' pNp and NMPs as products. To address this, Rrp44 decay products were analyzed by ion-pair reverse-phase HPLC. To validate our HPLC protocol, we first analyzed products of 3' phosphate RNA decay by the Rrp44 homolog RNase R from *E. coli* and observed quantitative 3'-5' pAp release, as previously observed using radiolabeled substrates and thin-layer chromatography (Cheng and Deutscher, 2002; Vincent and Deutscher, 2006) (Figures 24C through 24D). Quantitative release of 3'-5' pAp was also observed as a product of Rrp44 degradation for 3' phosphorylated but not 3' OH RNA substrates (Figures 22E, 24A and 24B), indicating that cleavage by Rrp44 occurs at the 5' phosphate of the terminal ribonucleotide, consistent with our structural model.

Figure 24. Analysis of Rrp44 activity on 3' phosphate RNA. (A through D) Ion-pair reverse phase HPLC analysis of RNA decay products. Reactions contained either Rrp44 (A and B) or *E. coli* RNAse R (C and D) and the indicated substrates. UV absorbance was monitored at 260 nm for detection of nucleotides (A and C) or 490 nm for detection of the 5' fluorescein label (B and D). (E) Multiple sequence alignment using Clustal Omega (Sievers et al., 2011) of Rrp44 (Dis3) from six organisms as well as the related enzymes Dis3l and Dis3l2 from human and RNase II and RNase R from *E. coli*. (F) Binding affinities of wild type and mutant Rrp44 for AU-rich RNA as measured by fluorescence anisotropy. Error bars indicate ±1 standard deviation obtained from three independent experiments. (G) RNA decay of 3' phosphate 5' FAM 49 nt AU-rich RNA by the indicated enzymes.



E

```
E.coliII          lfakalpddklqlivaiadptawiaegskldkaakiraftnylpgfNipmlprelsddlc   271
E.coliR           vycekkrgggwrlwvaiadvsyyvrpptpldrearnrgtsvyfpsqVipmlpevlsnglc   342
H.sapiensDis3L2   lsckpladgnfkvgvhiadvsyfvpegsdldkvaaeratsvylvqkVvpmlprllceelc   453
H.sapiensDis3L    lsvrtlnngnlelgvhiadvthfvapnsyidieartrattyyladrRydmlpsvlsadlc   548
D.rerio           lhcrelengnlevgvhiadvshfirpgnaldqeaanrgttvylcgkRidmvpellssnlc   546
X.laevis          lhcrelengtlevgvhiadvshfirpgnaldqeaanrgttvylcdkRidmvpellssnlc   544
H.sapiens         lhcrelengnlevgvhiadvshfirpgnaldqesarrgttvylcekRidmvpellssnlc   549
M.musculus        lhcrelsngnlevgvhiadvshfirpgnaldqesarrgttvylcekRidmvpellssnlc   549
S.cerevisiae      LHAKKLPNGNWEVGVHIADVTHFVKPGTALDAEGAARGTSVYLVDKRIDMLPMLLGTDLC   613
S.Pombe           lhacvlpngnyevgvhiadvthfvkpntsmdseaasrgttvylvdkRidmlpmllgtdlc   578
                  :          .: * *** : ::   . :*  .  *. : *:     *:*  *   **
                                                          R600
```



An attempt to disrupt recognition of 3' phosphate RNA identified Rrp44 Arg600 and Asp602, as they are conserved in eukaryotic homologs (Figure 24E) and located near the sulfate ion and 3' end of the RNA (Figure 22A). Differential binding defects were observed for the R600D/D602R mutant with 3' OH RNA preferred over 3' phosphate RNA (Figure 24F), however mutation of Arg600, alone or in combination with mutation of Asp602, resulted in diminished decay activities on both 3' phosphate and 3' OH RNA substrates (Figure 24G). This is perhaps consistent with amino acid substitutions at these positions in RNaseR and RNaseII as both enzymes are able to degrade 3' phosphorylated RNA substrates (Cheng and Deutscher, 2002) (Figure 24E). It is notable that Rrp44 homologs from human, including Dis3, Dis3l and Disl2, can also degrade 3' phosphate RNA by releasing pNp, suggesting that this activity may be evolutionarily conserved (Lubas et al., 2013; Tomecki et al., 2010). Although we were unsuccessful identifying separation of function mutations to selectively diminish Rrp44 activities on 3' OH or 3' modified RNA substrates, future efforts along this line will be critical to determine if Rrp44 participates in the degradation of 3' modified RNA substrates in vivo.

RNA path to Rrp6

The structure of Exo10$^{6exo\text{-}\Delta N\Delta C}$ bound to poly(A) RNA (PDB 4OO1) revealed 6-8 nt of the 24 nt substrate with the RNA 3' end anchored in the Rrp6 active site, the 5' end extending toward the S1/KH ring, and the remaining RNA disordered (Wasmuth et al., 2014). An alternative model was proposed based on a twinned crystal structure of the isolated Rrp6 catalytic domain bound to 15 nt polyU RNA modeled with partial occupancy (PDB 5C0Y) suggesting that RNA would not contact the S1/KH ring, and that an 18 nt AU-rich RNA bypassed the Rrp6 active site altogether in a nuclear exosome complex (Makino et al., 2015) (Figure 23).

Electron densities are evident for 17 nt of RNA in our structure up to the position where the di-triazole linker connects the two segments in our symmetric 34 nt RNA substrate (Figures 25A and 9D). As observed in the Exo10$^{6exo-\Delta N\Delta C}$ structure bound to poly(A) RNA (Wasmuth et al., 2014), the RNA 3' end is anchored in the Rrp6 active site, extending toward and contacting the S1/KH ring (Figures 9C and 25A). After reaching the bottom of the S1/KH ring, the RNA is deflected and makes a U-turn back toward the Rrp6 HRDC domain. Although not modeled, additional electron densities ~20 Å from the 5' end are observed between the HRDC and Exo domains of Rrp6 (Figures 9D, 25A and 25B), consistent with, but not identical to, a path proposed based on the twinned crystal structure of the Rrp6 catalytic domain bound to a 15 nt polyU RNA (Makino et al., 2015). Electron densities were not evident for the linker atoms, and RNA was not modeled in the HRDC/Exo domain channel due to poor definition of electron densities in this region.

As anticipated for a non-specific exoribonuclease, protein-RNA interactions are dominated by contacts to the polyribonucleotide backbone and non-specific stacking with the nucleobases of the RNA chain. Contacts between the last 3 nt of RNA (A_{17-15}) and Rrp6 are similar to those described for Exo10$^{6exo-\Delta N\Delta C}$ bound to poly(A) RNA (Wasmuth et al., 2014), however the paths diverge at the fourth nt (Figure 23). Instead, the nucleobase of A_{14} stacks on Rrp40 Arg110 while the A14 phosphate is coordinated by Rrp6 Tyr315 and His326 (Figure 25C). The U13 nucleobase is stacked on Rrp4 Phe177 and directed toward Rrp4 Ser175 (Figure 25D). The U12 and U11 nucleobases form base edge interactions with A2 and U4, respectively, with the U12 phosphate proximal to the Rrp41 Arg119 side chain and the U11 phosphate contacting Rrp4 Arg123. Density for the U10 nucleobase is not evident, but it points toward Rrp41 Lys62 (Figure 25D). The A9 nucleobase is sandwiched between Rrp40 Phe80 and the nucleobase of U7 while the A9 phosphate

contacts Rrp40 Lys85 (Figure 25E). The U8 nucleobase contacts Rrp40 Ser94 above Glu131 before turning to enable contacts between the U8 phosphate and Csl4 Arg150, a side chain that is also in proximity to the U7 phosphate (Figure 25E). The U6 nucleobase is stacked between A9 and A5 and makes contacts to the Rrp40 Ser81 side chain and backbone amide (Figure 25E). The A5 nucleobase stacks between U6 and U11 with its base edge directed toward A14 and the U4 nucleobase stacks between A5 and the A2 ribose with it base edge directed toward U12. The path twists to point the U3 nucleobase toward that of A16, with the A2 nucleobase contacting Rrp4 Met137 with its edge directed toward U12. The final U1 nucleobase is sandwiched between Rrp4 Leu138 and Rrp6 Tyr430.

Previous mutational analysis reported that Rrp6 Phe294 and Tyr315 did not play a role in Rrp6 mediated decay (Makino et al., 2015), although these residues contact RNA in both the Exo10$^{6exo\text{-}\Delta N\Delta C}$ and present structure (Wasmuth et al., 2014). Indeed, F294A/Y315A diminished Rrp6 activities when associated with Exo9 (Exo10$^{6\Delta N}$) or in the nuclear exosome (Figures 25F, 26A and 26C). Consistent with a path that guides RNA between the HRDC and Exo domains of Rrp6, Y244A and R245A substitution in Rrp6 reduced degradation activity in the exosome and as a free enzyme as noted previously (Makino et al., 2015) (Figures 25F, 26A and 26C). Rrp6 Arg461 and Asn466 are just below Tyr244 and Arg245, and R461V and N466D substitutions resulted in altered degradation patterns for poly(A) RNA when associated with Exo9 or when combined with Rrp40 K85E (see below; Figures 26B and 26D).

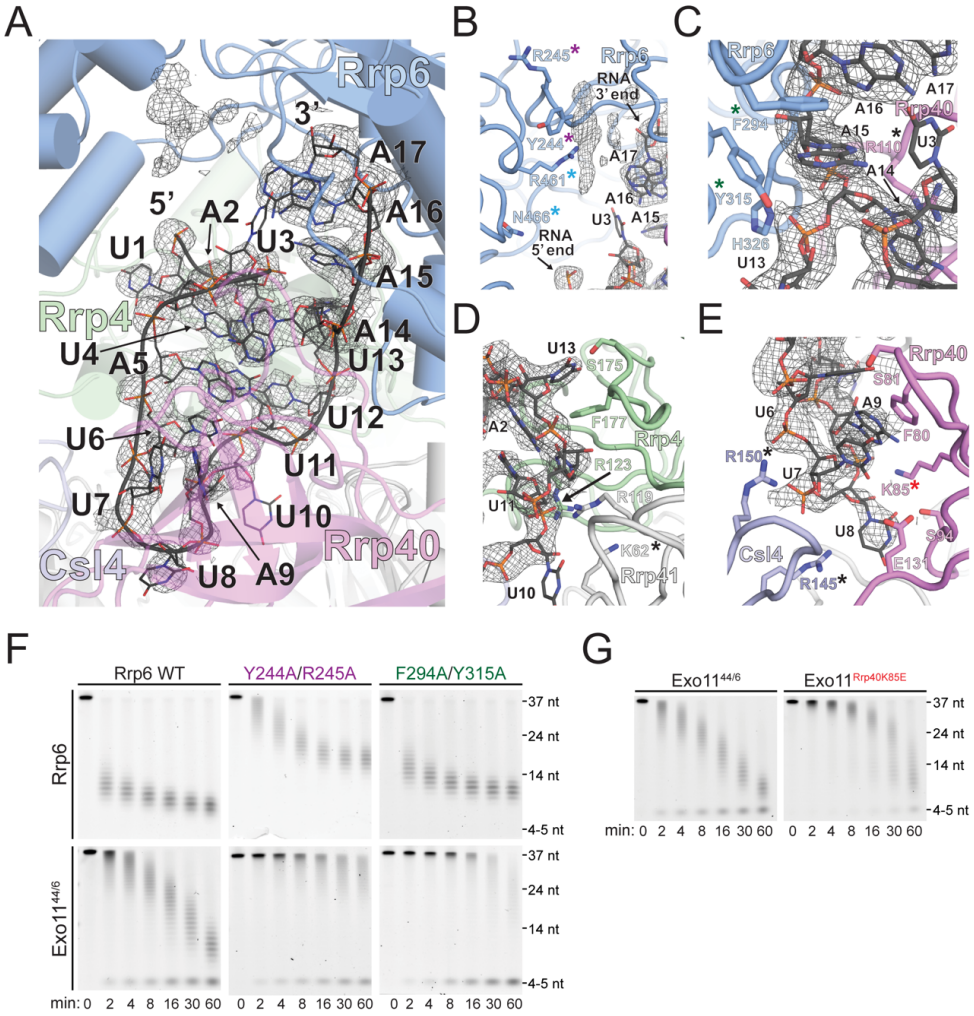


Figure 25. RNA path to Rrp6 in Exo11$^{44/6}$/3'-3'RNA structure. (A to E) RNA contacts to Rrp6, the S1/KH proteins, and Rrp41. Mesh represents electron density from a simulated annealing 2Fo-Fc omit map contoured at 1.2 σ. Unmodeled density between the Exo and HRDC domains of Rrp6 is shown (B). Residues mutated in this and/or previous studies that contribute to the activities of the RNA exosome are highlighted with colored asterisks. (F) Polyacrylamide TBE-urea gel analysis of decay of 5' FAM 37 nt poly(A) RNA by mutants in Rrp6. (G) Polyacrylamide TBE-urea gel analysis of decay of 5' FAM 37 nt poly(A) RNA by WT Exo11$^{44/6}$ and a mutant in the S1/KH protein Rrp40.

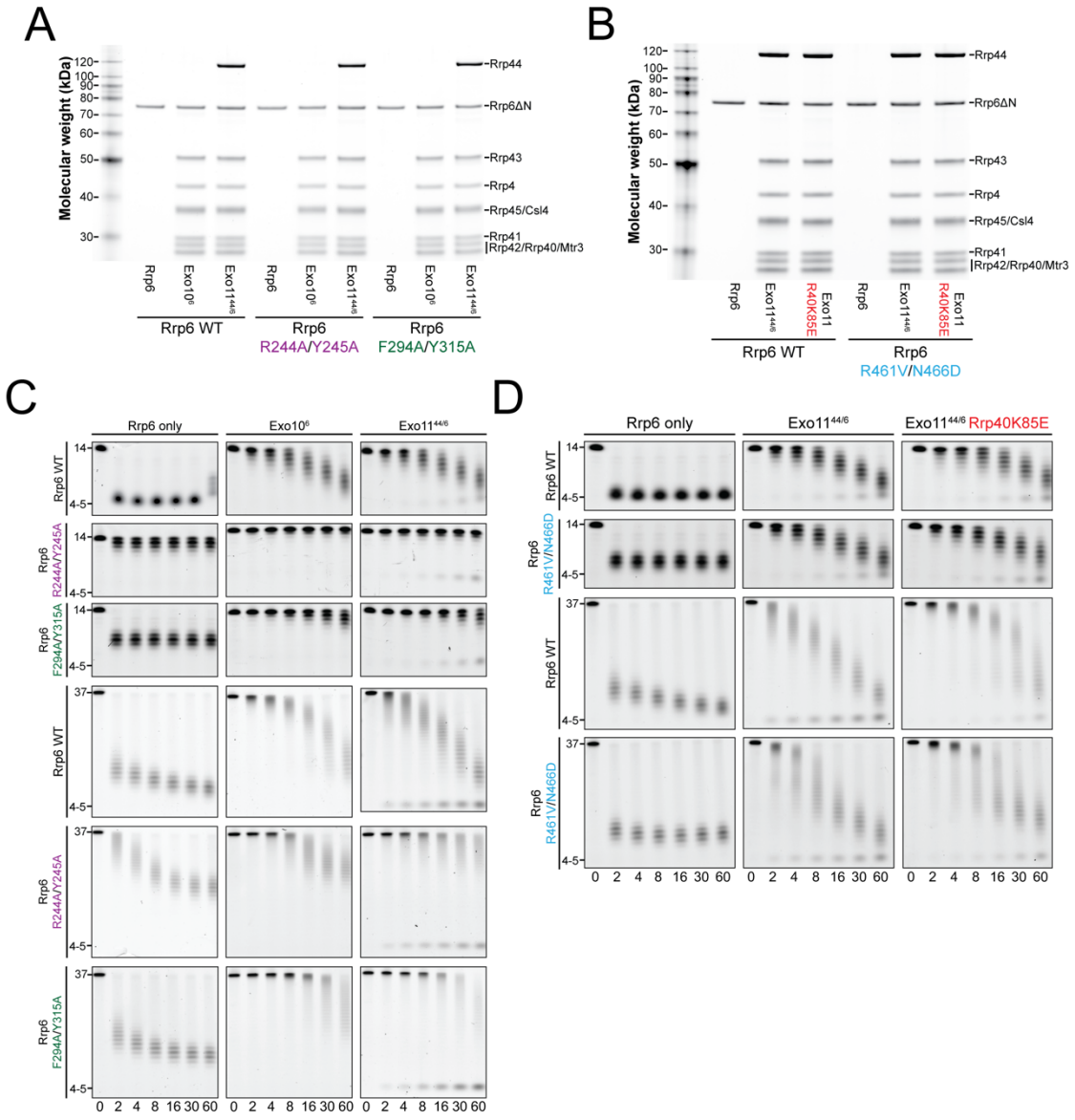


Figure 26. Rrp6 activity is diminished by mutating non-catalytic residues in its Exo domain and both Rrp6 and Rrp44 activities are diminished by a mutation in the S1/KH protein Rrp40. (A and B) SDS-PAGE analysis of Rrp6 wild type and mutant enzymes and reconstituted complexes. Gel is 4-12% acrylamide BIS-TRIS run in MES-SDS running buffer and stained with SYPRO ruby. (C and D) Polyacrylamide TBE-urea analysis of decay time courses of free Rrp6 and reconstituted complexes on 5' FAM poly(A) 14 nt and 37 nt RNA.

RNA contacts occur deep in the S1/KH ring with Csl4 Arg145 and Arg150 side chains contacting the phosphate backbone at U7 and U8, while Rrp40 Arg110 buttresses the A14 nucleobase (Figures 25C). Consistent with these contacts, prior analysis revealed that mutation of these conserved residues diminishes both Rrp6 and Rrp44 activities (Wasmuth et al., 2014). In addition, insertions between residues 63 and 64 of Rrp41 also disrupt Rrp6 and Rrp44 activity (Wasmuth and Lima, 2012a), consistent with occluding the RNA path to Rrp6 near U10 (Figure 25D). Among several unique contacts observed in our structure are those at the deepest point of RNA ingress to the S1/KH ring between the conserved Lys85 of Rrp40 (Liu et al., 2006) and the RNA backbone phosphate of A9 (Figures 25A and 25E). Of note, Lys85 lies just below that last nucleotide modeled in the channel-dependent RNA bound exosome structure (Makino et al., 2013b). Consistent with contacts observed in our structure and predicted contacts in the channel-dependent structure, Rrp40 K85E decreased both Rrp6 and Rrp44 activities, supporting a role in coordinating RNA via overlapping but distinct paths as RNA is guided to Rrp6 or Rrp44 (Wasmuth and Lima, 2012a; Wasmuth et al., 2014) (Figures 25G, 26B and 26D).

Discussion

The strategy employed here to satisfy both Rrp6 and Rrp44 exoribonuclease sites with a single RNA containing two 3' ends enabled determination of a high-resolution structure of the nuclear RNA exosome, and it may prove useful in characterizing other RNA exosome complexes in association with additional nuclear cofactors.

Association of Rrp6 and Rrp44 with the essential non-catalytic Exo9 core attenuates or alters their activities (Dziembowski et al., 2007; Wasmuth and Lima,

2012a). Previous studies suggested that attenuation is due in part to sequestration of their active sites relative to the apo enzymes by imposing a requirement that RNA must pass into the S1/KH ring to be degraded by Rrp6 or through the entire central channel to be degraded by Rrp44 (Drazkowska et al., 2013; Makino et al., 2013a; 2013b; Malet et al., 2010; Wasmuth and Lima, 2012a; Wasmuth et al., 2014). Using the high-resolution structure presented here, extensive interactions were revealed between the RNA and S1/KH ring of central channel on its way to the Rrp6 active site. Mutational analysis presented here and elsewhere suggests that amino acid sidechains along this path can contribute to both Rrp6 and Rrp44 activities, supporting the existence of overlapping yet distinct RNA paths that guide substrates to the respective catalytic subunits (Makino et al., 2015; Wasmuth and Lima, 2012a; Wasmuth et al., 2014) (Figure 11A). We also resolved a feature of the Exo9 core, namely Rrp43L1, which together with $Rrp45_{Cterm}$ interact with the channel-independent Rrp44 conformation (Figures 11A, 11C, and 12A through 12C) and inhibit Rrp44 activity. We found that deleting these features enhanced Rrp44 activity on long and short RNAs, and that activities remained dependent on the Exo9 central channel, especially for longer RNA substrates (Figures 14 through 16).

One alternative explanation for enhanced activity in the absence of Rrp43L1 is that Rrp43L1 is highly acidic and contacts Rrp44 near the RNA point of ingress. As such, it might create an electrostatic barrier that inhibits binding to the direct-access or channel-independent Rrp44 conformation. However, this model predicts that binding and decay of RNA by the Rrp43L1 deletion mutant should not be affected by a channel occlusion nor should its activities be suppressed by deletion of Rrp43L2, which is not the case (Figures 14 through 16). We propose a model in which Exo9 elements that bind and stabilize the channel-independent Rrp44 conformation (Rrp43L1 and $Rrp45_{Cterm}$) compete with Exo9 elements that bind and stabilize the

channel-dependent Rrp44 conformation (Rrp43L2). Since deletion of Rrp43L2 suppresses the activities gained by deleting Rrp43L1 and $Rrp45_{Cterm}$ on both long and short RNA substrates, we propose that conformational changes to the channel-dependent Rrp44 conformation are required for the most efficient degradation of these substrates.

A recent genetics study in *S. cerevisiae* showed that a strain expressing only Rrp44da (Figure 17) accumulated certain RNA species such as $tRNA_i^{Met}$ and the rRNA 5' externally transcribed spacer (5' ETS) (Han and van Hoof, 2016). Intriguingly, the authors showed that expressing this mutant Rrp44 rescued growth defects but not 3' processing defects in strains containing channel occlusions in the Rrp45 and Rrp41 proteins (Figures 18B and 18C). However, these results are complicated by our finding that this mutant contains an intrinsic loss of activity of the enzyme (Figure 19C), which could contribute to the accumulation of these substrates. Furthermore, a loss of exo activity could explain the synthetic effect with combining the da mutant with the endo-. Another genetics study purporting to selectively disrupt the channel-independent Rrp44 conformation employed a mutant in the S1 domain that we would predict to disrupt the free enzyme activity as well (Delan-Forino et al., 2017). While the Rrp45 and Rrp43 loop deletion mutants we employ do alter Rrp44 activity in the RNA exosome complex (its presumed physiological state), their use offers the advantage that they are not alterations in Rrp44 itself, and thus cannot affect its intrinsic activity. Use of these mutants in future genetics studies may allow for unambiguous determination of the effects of perturbing Rrp44 conformational switching in vivo.

The models presented here and elsewhere provide a structural basis for the ability of Rrp44, but not Rrp6, to degrade 3' phosphorylated RNA substrates. The in vivo significance of these findings are discussed extensively in the Introduction

section of this dissertation. Previous studies implicated Xrn1 in degradation of 2',3'-cyclic phosphate tRNA introns (Wu and Hopper, 2014), but the observation that Rrp44 and related human enzymes can degrade 3' phosphorylated RNA begs the question as to whether the RNA exosome also participates in the degradation of aberrant 3' modified forms of RNA, or if this activity contributes to generating clean 3'OH RNA ends to promote post-transcriptional polyadenylation by TRAMP (LaCava et al., 2005; Wyers et al., 2005) or 3' trimming by Rrp6.

Materials and Methods

Cloning, expression, and purification of proteins

S. cerevisiae Exo9 components were cloned as N-terminal His_6Smt3 fusions, expressed in *E. coli*, and reconstituted into complexes as previously described (Greimann and Lima, 2008; Wasmuth and Lima, 2012a). Briefly, Smt3 was amplified from *S. cerevisiae* genomic DNA using PCR primers that added six histidines and a short linker to its N-terminus and cloned into the NcoI and BamHI sites of pRSF-duet (Novagen) to create pRSFduet-Smt3. PH-like subunits were cloned into this vector as heterodimers as follows: His_6Smt3-Rrp42 paired with Mtr3, His_6Smt3-Rrp46 paired with Rrp43, and His_6Smt3-Rrp41 paired with Rrp45. Cap subunits Csl4, Rrp4 and Rrp40 were cloned into the first multiple cloning site of pRSFduet-Smt3 and expressed alone. Plasmids were transformed into BL21 (DE3) RIL cells (Agilent) and single transformants were inoculated into antibiotic containing Superbroth (Teknova) and incubated with shaking at 37°C. Cultures were induced at OD_{600} = 1.0 – 2.0 with cold shocking on slushy ice and 400 µM IPTG, grown at 18°C with shaking overnight and harvested by centrifugation. Cell pellets were resuspended in an equal volume of 50 mM Tris-Cl pH 8.0, 20% w/v sucrose and stored at -80°C. Prior to lysis, the suspension was thawed and adjusted to 350 mM NaCl, 20 mM imidazole, 1 mM beta-mercapto ethanol (BME), 0.1 % v/v IGEPAL co-630, 10 µg/mL lysozyme and 10 µg/mL DNAse I. Cells were lysed by sonication with stirring on ice. The lysate was separated from unlysed cells and insoluble matter by centrifugation at 40,000 x g (4°C) for 45 min, added to Ni-NTA resin (Qiagen), and rotated for 1 hr at 4°C. Then, the resin and lysate were transferred to a disposable column, the flowthrough discarded, and the resin washed with 10 volumes of wash buffer (20 mM Tris-Cl pH 8.0, 250 mM NaCl, 20 mM imidazole, 1 mM BME) using gravity flow. The resin was

then washed with 10 volumes of chaperone buffer (wash buffer plus 50 mM KCl, 10 mM $MgSO_4$, and 2 mM ATP) followed by another 5 volumes of wash buffer. Bound protein was eluted with 3 column volumes of wash buffer plus 250 mM imidazole pH 8.0 and injected onto a 320 mL Superdex200 (for the PH-like protein heterodimers and Rrp4) or Superdex75 columns (for Csl4 and Rrp40) that had been equilibrated in 20 mM Tris-Cl pH 8.0, 350 mM NaCl, 1 mM NaCl. Peak fractions were pooled, concentrated to >5 mg/mL, and a 1:1000 molar ratio of Ulp was added and incubated at 4°C with the protein or heterodimer overnight to cleave the tag. The next morning the preparation was flash frozen and stored at -80°C for later use. For the Rrp46/43 heterodimer, the Ulp cleavage step was omitted and the preparation was simply concentrated and flash frozen after Superdex elution. For Exo9 reconstitution, purified protein stocks were thawed, centrifuged, and the concentrations of the supernatants measured using a Bradford assay. Then, they were mixed in a molar ratio of Rrp41/Rrp45 : Rrp42/Mtr3 : Rrp46/Rrp43 : Csl4 : Rrp40 : Rrp4 = 1 : 1 : 3 : 1.75 : 1.75 : 1 and dialyzed into 20 mM Tris-Cl pH 8.0, 100 mM NaCl, 0.1 mM $MgCl_2$, 10 mM DTT at 4°C overnight. The next day the dialysate was injected over a Superdex200 column that had been equilibrated in the dialysis buffer, peak fractions were pooled and concentrated to >5mg/mL, flash frozen in liquid nitrogen and stored at -80^{o}C for later use.

Genes for the Rrp43 deletion mutants were synthesized by Life Technologies and were as follows: Rrp43ΔL1 has residues 101-119 of Rrp43 replaced with Gly-Ser-Gly-Ser-Gly-Ser, Rrp43ΔL2 has residues 254-270 deleted, and Rrp43ΔΔ contains both of these modifications. Rrp45ΔCterm (residues 1-290) has the C-terminal 14 nt deleted and was generated using standard PCR/restriction enzyme based cloning. Point mutants in Rrp6, Rrp40, and Rrp44 were introduced by PCR based site-directed mutagenesis. All mutants were verified by Sanger sequencing of the entire gene. For

cloning of the Rrp6/Rrp47 heterodimer used in the complexes in Figures 19 to 21, Rrp47 was cloned into the first multiple cloning site of pRSFduet-Smt3 and full length Rrp6 was cloned into the second.

For reconstitution of complexes containing catalytic subunits, Exo9 cores were mixed with an equivalent of Csl4, a 3-fold molar excess of Rrp6 (or Rrp6/47 for Figures 19 through 21) and/or 1.5-fold molar excess of Rrp44, incubated at 4°C for 2-4 hrs, and separated from free enzymes using a Superdex200 increase 10/300 GL column (GE) in 20 mM Tris-Cl pH 8.0, 100 mM NaCl, 0.1 mM $MgCl_2$, 0.5 mM TCEP-HCl. All preparations of Rrp6 in this chapter except for in the experiments presented in Figures 19 to 21 lack the N-terminal 128 residues, which form an obligate heterodimer with the cofactor Rrp47 (Feigenbutz et al., 2013b; Schuch et al., 2014). Rrp6 containing the full C-terminus was expressed as an N-terminal His6-Smt3 fusion in *E. coli* BL21 (DE3) RIL cells with 250 µM IPTG induction for 16 hours at 18°C. The next steps are exactly the same as for the core subunits up to the nickel elution, at which point the Nickel eluate is injected onto a 5 mL Hi-Trap Heparin HP column (GE) equilibrated in 20 mM Tris-Cl pH 8.0, 250 mM NaCl, 1 mM BME. The column was washed with 4 column volumes of the equilibration buffer and eluted with a linear gradient of NaCl to 1 M over 20 column volumes. Rrp6 eluted in a peak centered at approximately 430 mM NaCl. Fractions containing the Rrp6 peak were pooled, concentrated, and digested with 1:1000 molar ratio of SUMO protease overnight at 4°C. The next day, the protein was separated from His_6Smt3 and SUMO protease using a Hiload 26/600 Superdex 200 (GE) gel filtration column (in 20 mM Tris-Cl pH 8.0, 300 mM NaCl, 2 mM BME). Fractions containing Rrp6 were pooled, concentrated to 7-10 mg/mL by centrifugation in an Amicon YM-30 filtration unit (Millipore), flash frozen in liquid nitrogen, and stored at -80°C. This protocol effectively rids the preparation of nucleic acid contamination and C-terminal truncated

products. The same protocol was used to purify *E. coli* RNAse R (cloned as an N-terminal His6-Smt3 fusion), except the gradient for the heparin column was run from 250-600 mM KCl in 20 mM Tris-Cl pH 8.0, 0.1 mM $MgCl_2$, 5% v/v glycerol, 0.5 mM TCEP-HCl and the protein eluted in a peak at approximately 450 mM KCl. Expression and purification of the Rrp6/Rrp47 heterodimer was nearly the same as for Rrp6ΔN, except for that it was transformed into SoluBL21 (DE3) RIL cells (Genlantis) that were grown and induced in a 12 L fermenter rather than flasks. Additionally, the protein elutes at higher NaCl concentrations (~600 mM) from the heparin column and its Smt3 tag remains uncleaved until reconstitution (see Chapter 2). Heparin purification for Rrp44 is detailed in the next chapter.

Synthesis and HPLC purification of 3'-3' RNAs

5' alkynyl RNA oligonucleotides were synthesized and purified by standard desalting by Integrated DNA technologies. 17 and 18 sequences were 5' (hexynyl)UA UUA UUU AUU UUA AAA 3' and 5' (hexynyl)UUA UUA UUU AUU UUA AAA 3', respectively. Copper click chemistry was carried out via modification of published protocols (El-Sagheer and Brown, 2012; Paredes and Das, 2011). Click reactions (10-200 µL) contained 200 µM 5' hexynyl oligonucleotide, 5 mM sodium ascorbate, 80 mM potassium phosphate pH 7.0 and 100 µM 1,5 diazido-3-oxapentane (Santa Cruz Biotechnology, 10 mM stock made in DMSO). Reactions were initiated with the addition 1/10 volume of 5 mM $CuCl_2$, 5 mM Tris(3-hydroxypropltriazolemethyl) amine (Sigma-Aldrich) and incubated at 25°C for 30 minutes. Reactions were quenched by addition of EDTA to 10 mM, diluted to 500 µL with TE buffer and fractionated by DEAE chromatography as described below. The concentration series for the 1,5 diazido-3-oxapentane titration shown in Figure 9B is 10, 25, 50, 100, 200, 400, and 1000 µM. Click reactions were fractionated by DEAE chromatography

(Waters Protein-Pak DEAE 8HR 1000 Å 8 μm 5x50 mm column) using an Alliance 2695 HPLC separation module (Waters). The column was run at a flow-rate of 0.5 mL/min at 50°C and equilibrated in buffer A (10 mM Tris-Cl pH 8.0, 400 mM NaCl, 0.1 mM EDTA). A linear gradient to 80% buffer B (10 mM Tris-Cl pH 8.0, 600 mM NaCl, 0.1 mM EDTA) was run from 2.5 min to 30 min after injection and fractions were collected every 1 minute (Figures 10A and 10B). Fractions were analyzed by UREA-PAGE using SYBR Gold staining (Life Technologies), and those containing the desired product were adjusted to 300 mM sodium acetate pH 5.2 and 70% ethanol to precipitate the RNA. The precipitate was then pelleted by centrifugation at 20,000 x g (4°C), the solvent decanted, the pellet washed with cold 70% ethanol, centrifuged again, the solvent decanted, and the pellet air-dried for 1 hr and resuspended in approximately $1/10^{th}$ the original reaction volume of RNAse free water (Ambion). RNA concentration was calculated by measuring A_{260} using a Nanodrop 2000 and dividing that value by a theoretical extinction coefficient based on its sequence. This procedure regularly yielded 40-55% of the expected 3'-3' RNA product at concentrations of 400-550 μM.

Crystallization of Exo11 in complex with 3'-3' RNA

Exo11$^{Rrp44\ D171N/D551N,\ Rrp6(residues129\text{-}684)D238N}$ at a concentration of 13-15 mg/mL (by Bradford assay) in 20 mM Tris-Cl pH 8.0, 100 mM NaCl, 0.1 mM $MgCl_2$, 1 mM TCEP-HCl was mixed in a 1.0:0.9 protein:RNA molar ratio with 3'-3' RNA and incubated for 2 hours on ice prior to mixing with well solution. Then, 0.8 μL of protein/RNA solution was mixed with 0.4 μL of well solution (100 mM sodium citrate pH 5.25, 7 mM NaMES pH 6.5, 6.5-7% PEG 3350, 175 mM $(NH_4)_2SO_4$, 0.25-1 mM $MgCl_2$; final pH 5.6) in 24-well hanging drop format, sealed and incubated at 4°C. Crystals appeared after 1.5 days and grew to their full size (~350 μm in the largest

dimension) after 1-2 weeks. Prior to freezing in liquid nitrogen, crystals were transferred to well solution containing 5% v/v glycerol, incubated for 1-5 min, then transferred in four steps with increasing glycerol to final well solution containing 30% v/v glycerol.

Structure determination.

X-ray diffraction data were collected at the Advanced Photon Source 24-ID-C beam line equipped with a Pilatus-6MF detector. Data was obtained from a single crystal diffracted at a wavelength of 0.9795 Å at 100 K and processed using HKL2000 (Otwinowski and Minor, 1997). Statistics reported in Table 1 were obtained using Phenix (Adams et al., 2010). The structure was solved by molecular replacement using Phaser (McCoy et al., 2007) and coordinates of the yeast Exo10[6] (PDB: 4OO1) and yeast Rrp44 (PDB: 2WP8) as search models. The structure is refined to R/R_{free} values of 0.249/0.201. The final model includes 24 of the 34 possible nucleotides, and 3731 of the 4158 amino acids present in the crystal. Analysis of the contents of dissolved crystals is consistent with 1:1 RNA:exosome complex (Figures 10E and 10F). The asymmetric unit contains one complex. Iterative rounds of refinement were accomplished using Phenix (Adams et al., 2010). RNA and side chains were manually built using Coot (Emsley et al., 2010). Simulated annealing omit maps and maps used during building were also generated using CNS (Brunger, 2007). The model was refined using positional refinement, real-space refinement and individual B-factor refinement. Figures depicting the structure, including the surface representations in Figure 2A and the graphical abstract, were prepared with Pymol (Schrödinger). Structure quality was assessed using MolProbity (Chen et al., 2010) indicating the model has excellent geometry with 94.6% in favored, 5.2% in allowed, and 0.2% in

outlier regions of Ramachandran space. The structure scored in the 100th percentile for the Clash and MolProbity scores.

RNA binding and decay assays

AU-rich RNA sequences (Integrated DNA Technologies) were A AUU AUU UAU UAU UUA UUU AUU AUU UAU UUA UUU AUU AUU UAU UUA UUA, AUU AUU UAU UUA UUA AUU AUU UAU AUU UUA UUU AUU, and AU UUA UUU AUU AUU for the 49, 36 and 14 nt RNA and were HPLC purified by Integrated DNA technologies. The sequence of the 5'-3' RNA used in Figure S1C is AAA AUU UUA UUU AUU UAU UUA UUA UUU AUU UUA AAA to mimic that of the corresponding 3'-3' RNA. PolyA RNAs (Dharmacon) were of the sequence A_x, where x = length in nucleotides of the indicated species, and were HPLC purified by Dharmacon. Unless otherwise noted (for example in the experiments in Figures 19 and 20), decay assays contained 1 nM enzyme and 10 nM RNA substrate in RNA decay buffer (20 mM Tris-Cl pH 8.0, 50 mM KCl, 0.5 mM $MgCl_2$, 0.5 mM TCEP-HCl, 0.1 % IGEPAL co-630). For gel analysis of decay products, an aliquot of decay reactions containing 5' fluorescein or 5' 6-carboxyfluorescein (FAM) RNAs was quenched at the indicated times with an equal volume of 89 mM Tris-borate pH 8.3, 7 M Urea, 2 mM EDTA, 12% w/v Ficoll, 0.005 % w/v xylene cyanol and 10 μL was loaded onto 15% acrylamide TBE urea gels (Life Technologies). Gels were imaged for fluorescein fluorescence using a Typhoon FLA9500 instrument (GE). For measuring initial rates of decay of Exo10^{44} or Rrp44 in Exo11$^{44/6}$ (Figures 14D and 15D), the Rrp44 product was quantified at the first two to three non-zero time points in triplicate using ImageJ (NIH) and converted to nM via a standard curve of fully decayed RNA. Initial rate is calculated as the slope of the linear regime of decay under these conditions. For quantification of Rrp44 activity in Exo10^{44} and Exo11$^{44/6}$ complexes (Figures 14B,

14C, 15B, 15C, 15E and 15F), nM of Rrp44 product was calculated in the same fashion. For decay assays using 3' phosphate RNAs (Figures 22D and 24G), reactions were incubated at 30°C for 5 minutes in the presence or absence of 0.01 U/μL calf intestinal phosphatase (CIP, from New England Biolabs) prior to addition of the exonucleases as indicated. For K_D measurements, binding reactions (20 μL) contained 30 nM 5' FAM RNA and variable protein concentrations in RNA decay buffer. Proteins were buffer exchanged into RNA decay buffer prior to the binding experiment using Micro Biospin 6 columns (Bio-Rad). To render them catalytically inactive, Rrp44 contained D171N and D551N mutations and $Rrp6_{129-733}$ contained a D238N mutation. Reactions were incubated for 20 minutes at room temperature in black low-volume, non-binding, round-bottom 384 well plates (Corning) prior to anisotropy measurement (λ_{ex} = 495 nm, λ_{em} = 525 nm) using a Spectramax M5 plate reader (Molecular Devices). K_Ds were calculated from a triplicate measurement using eleven 1:2 serial dilutions of a given protein plus a -protein control using a receptor depletion model to fit (Wasmuth and Lima, 2012a).

Ion-pair reverse phase HPLC analysis of decay products

20 μL RNA decay reactions containing 1 μM enzyme and 10 μM 49 nt 5' FAM AU-rich RNA (or mock substrates) were incubated for 25 minutes at 30°C in RNA decay buffer lacking IGEPAL CO-630. After incubation, reactions were quenched with the addition of EDTA to 10 mM, adjusted to 10 mM $N(C_4H_9)_4HSO_4$ (Sigma), and the products separated by HPLC using a Nova-Pak C18 60 Å 4 μM 3.9 x 150 mm reverse-phase column (Waters). The column was equilibrated with RP buffer A (50 mM potassium phosphate pH 7.0, 10 mM $N(C_4H_9)_4HSO_4$) and run at a flow rate of 1 mL/min at 40°C. Products were eluted with a linear gradient to 100% RP buffer B (RP buffer A containing 50% v/v acetonitrile) from t=5 to t=20 minutes (Figures 22E

and 24A through 24D). UV absorbance was monitored at 260 nm for detection of nucleotides and small RNA decay products (Figures 22E, 24A, and 24C) and 490 nm for detection of the 5' FAM label (Figures 24B and 24D).

CHAPTER 2: Crystal Structure of the Mpp6-Exosome and Reconstitution of Helicase-Dependent RNA Decay

Introduction

Biochemical and structural characterization of the RNA exosome complex has, with few exceptions (Kowalinski et al., 2016; Makino et al., 2015; Wasmuth et al., 2017; Falk et al., 2017), been performed in the absence of its protein cofactors. The cofactor-less studies revealed that Exo9 modulates the activities of Dis3 and Rrp6; that the Rrp6 protein can stimulate Dis3; that RNAs greater than 30 nucleotides utilize the Exo9 central channel to guide RNA to Rrp6 for distributive trimming or through the entire channel to Dis3 for processive decay; and the existence of a minor channel-independent "direct-access" route to Dis3 for shorter RNAs (Liu et al., 2006; Makino et al., 2013b; Wasmuth and Lima, 2012a; Wasmuth et al., 2014; Zinder et al., 2016). However, genetic studies in diverse model systems suggest that most of the important functions of the RNA exosome are facilitated by its protein cofactors (Kilchert et al., 2016; Lee et al., 2013; Milligan et al., 2008; Mitchell et al., 2003; Schilders, 2005), prompting a need for rigorous biochemical and structural characterization of these complexes.

The nuclear exosome is associated with the cofactors Mpp6 and Rrp47, whose individual deletion in vivo results in misprocessing of nuclear exosome substrates, while concomitant loss leads to synthetic lethality (Butler and Mitchell, 2011; Callahan and Butler, 2008; Feigenbutz et al., 2013a; Milligan et al., 2008). These genetic observations resemble deletions of the TRAMP complex, a three protein complex that catalyzes polyadenylation and subsequent degradation of an array of nuclear RNAs (LaCava et al., 2005; Vanacova et al., 2005; Wyers et al., 2005)

through one of two non-templated poly(A) polymerases Trf4/5, one of two zinc knuckle proteins Air1/2, and the essential RNA helicase Mtr4. This complex is thought to bridge the nuclear exosome to other RNPs, including ribosome biogenesis factors (Thoms et al., 2015). Indeed, Rrp47 binds directly to Rrp6 via the Rrp6 N-terminal domain, an interaction that stabilizes Rrp6 and Rrp47 (Dedic et al., 2014; Feigenbutz et al., 2013a; 2013b; Stead et al., 2007). This interaction also generates a composite interface that recruits Mtr4 (Schuch et al., 2014) in vitro and in vivo, though other factors are speculated to exist (Stuparevic et al., 2013). In contrast, Mpp6 function is less clear, due in part to its poor sequence conservation. Its synthetic lethality with Rrp6 in yeast suggested Mpp6 may be a Dis3 cofactor (Milligan et al., 2008), while in human it was posited as a Rrp6 cofactor (Schilders, 2005) and has been reported to recruit Mtr4 (Chen et al., 2001; Lim et al., 2017; Lubas et al., 2011; Schilders et al., 2007).

To better understand how Mpp6 influences nuclear exosome activities and Mtr4 recruitment, we present the functional characterization of nuclear exosome complexes with Mpp6, Rrp47 and Mtr4 from the budding yeast *S. cerevisiae* and the crystal structure of a 12-subunit nuclear exosome complex bound to Mpp6 and RNA to a resolution of 3.3 Å. Biochemical and genetic experiments suggest that Mpp6 can bind RNA and stimulate the activities of the nuclear exosome, and that both Mpp6 and Rrp47 contribute to recruitment of the Mtr4 helicase to facilitate activation of the nuclear exosome and degradation of structured RNA.

Results

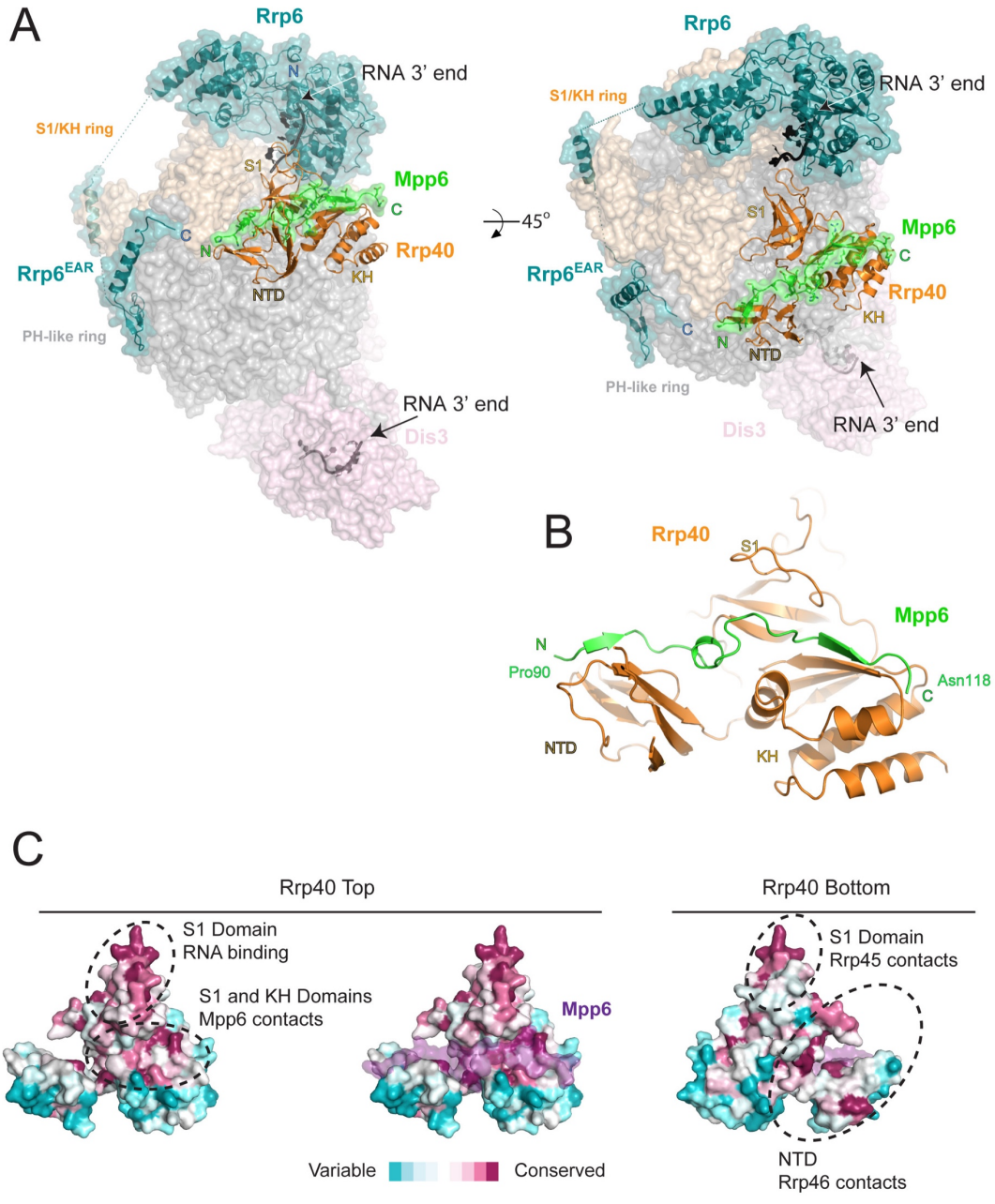
Mpp6 is anchored to Exo9 via contacts to the S1/KH ring subunit Rrp40

To facilitate crystallization of the nuclear exosome bound to a minimal fragment of Mpp6 (residues 81-120; $Mpp6_{minimal}$) and RNA, copper 'click chemistry' was used to generate a synthetic 3'-3' RNA that was designed to simultaneously engage both exonuclease sites of the complex as previously (Zinder et al., 2016) (Chapter 1). This RNA was incubated with the complex, and well solutions similar to those used to crystallize $Exo11^{Dis3exo\text{-}endo\text{-}/Rrp6exo\text{-}\Delta N\Delta C}$ were screened (Zinder et al., 2016). These efforts resulted in crystals which closely resembled Exo11 in morphology, final size, and growth kinetics. After screening several crystals, diffraction was collected to 3.3 Å and the structure was solved (Table 3, Figure 27A). Similar to the previous structure, one RNA 3' end is anchored in the Dis3 active site with Dis3 adopting a direct-access conformation (Liu et al., 2014; 2016; Makino et al., 2015; Zinder et al., 2016) while the other 3' end is anchored in the Rrp6 active site of an adjacent complex (Zinder et al., 2016) (Figure 9D). Although only 4-5 residues of RNA can be visualized in the Rrp6 active site, noncontiguous densities were observed proximal to residues previously identified to contact RNA, including Arg110 of Rrp40, and Rrp6 residues within the EXO and HRDC domains, in particular Tyr244, Tyr430, and Arg461 (Makino et al., 2015; Zinder et al., 2016). Densities corresponding to the loop and pair of helices in Rrp43 previously referred to as 'Rrp43L1' (Zinder et al., 2016) were not observed.

Table 3. Crystallographic data and refinement statistics. One crystal was used. Highest resolution shell is in parentheses. From Wasmuth et al., 2017.

Data collection	
X-ray Source	APS GM/CA 23IDD
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	141.1, 213.6, 225.9
α, β, γ (°)	90.0, 90.0, 90.0
Wavelength (Å)	1.0332
Resolution (Å)	44.3–3.3 (3.42–3.3) *
R_{merge}	0.086 (0.613)
$I/\sigma I$	9.1 (1.7)
$CC_{1/2}$	0.997 (0.161)
Completeness (%)	97.0 (95.0)
Redundancy	3.4 (2.6)
Wilson B factor ($Å^2$)	99.7
Refinement	
Resolution (Å)	44.3–3.3
No. reflections observed	341339
No. unique reflections	100440
R_{work}/R_{free}	0.217/0.266
No. atoms	29498
Protein	29147
RNA	249
Ligands	38
Water	64
Average B-factors	
Protein	138
RNA	139
Ligands	141
Water	61
R.m.s deviations	
Bond lengths (Å)	0.001
Bond angles (°)	0.41
Ramachandran plot	
% favored	93.5
% allowed	6.5
% outliers	0
Molprobity	
Clashscore/Percentile	$5.44/100^{th}$
MolProbity Score/Percentile	$1.72/100^{th}$

Figure 27. Structure of the 12-subunit Mpp6 nuclear exosome. (A) Global view of Exo12$^{Dis3exo\text{-}endo\text{-}/Rrp6exo\text{-}/Mpp6Min}$ bound to a 3'-3' RNA. Mpp6^{Minimal} interacts with an extended surface across the S1/KH subunit, Rrp40. View from side (left) and top (right). Exosome subunits shown as surface view, Mpp6 in green, Rrp6 in teal, Rrp40 in cartoon (orange), RNA in black sticks. (B) Mpp6^{Minimal} (green) makes extensive contacts to Rrp40 (orange) and spans all three of its domains. (C) Mpp6^{Minimal} (transparent purple surface in middle and right panels) binds to a conserved surface of Rrp40. Other conserved surfaces important for RNA binding and scaffolding interactions to other exosome subunits are indicated. Surface conservation calculated with ConSurf (Ashkenazy et al., 2010). Data acquired by Dr. Elizabeth Wasmuth with reagents prepared by John Zinder and Dr. Elizabeth Wasmuth. Figure prepared by Dr. Elizabeth Wasmuth with assistance from Dr. Christopher Lima. Figure and caption from Figure 2 of Wasmuth et al., 2017.


A
Rrp6
RNA 3' end
S1/KH ring
S1
Mpp6
C
N
Rrp40
Rrp6EAR
KH
NTD
PH-like ring
RNA 3' end
Dis3
45°
Rrp6
RNA 3' end
S1/KH ring
S1
Mpp6
C
Rrp6EAR
Rrp40
KH
C
N
NTD
PH-like ring
Dis3
RNA 3' end
B
Rrp40
S1
Mpp6
N
Pro90
Asn118
C
NTD
KH
C
Rrp40 Top
S1 Domain
RNA binding
S1 and KH Domains
Mpp6 contacts
Mpp6
Variable
Conserved
Rrp40 Bottom
S1 Domain
Rrp45 contacts
NTD
Rrp46 contacts

Electron densities for Mpp6 residues 90 to 118 were evident on the surface of Rrp40. $Mpp6^{Minimal}$ residues 90-97 adopt a distorted parallel beta-strand configuration that complements the second beta-strand of the Rrp40 N-terminal domain (NTD) while residues 98-103 adopt a helical conformation that wedges between the Rrp40 NTD and S1 domain (Figure 27B). Mpp6 residues 110-118 adopt an anti-parallel beta-strand conformation that complements the second beta-strand of the KH domain (Figure 27B). Although modeled, densities for amino acids 104-107 were very weak in comparison to other portions of Mpp6. Interactions between Rrp40 and Mpp6 are mediated by Rrp40 surfaces with comparable conservation to those involved in contacts to Rrp45 and Rrp46 or RNA (Malet et al., 2010; Wasmuth et al., 2014; Zinder et al., 2016) (Figure 27C). Contacts observed in our structure are also consistent with crosslinking and mass spectrometry data reported previously as well as a recent crystal structure of $Mpp6_{Minimal}$ bound to Exo9 (Falk et al., 2017; Shi et al., 2015) (Figure 29).

The most prominent electron density observed for Mpp6 corresponds to Arg112 within the last beta strand. Arg112 projects into a pocket formed between the S1 and KH domains of Rrp40 (Figure 28A) and resides in one of two regions of Mpp6 previously identified as conserved from yeast to man (Milligan et al., 2008). Given its location in the structure and evolutionary conservation, we refer to Mpp6 Arg112 as the “arginine anchor”. Mutation of this residue results in a loss of interaction of the exosome core in human (Falk et al., 2017) and yeast (Wasmuth et al., 2017) models. Interestingly, two known mutations within Rrp40 that are associated with neurological disorders lie proximal to the arginine anchor of Mpp6. These include Gly148 (Gly191 in human), which is mutated to cysteine in hereditary spastic paraplegia in humans (Halevy et al., 2014), and Trp195 (Trp238 in human), which is mutated to arginine in pontocerebeller hypoplasia and spinal motor neuron degeneration (Wan et al., 2012).

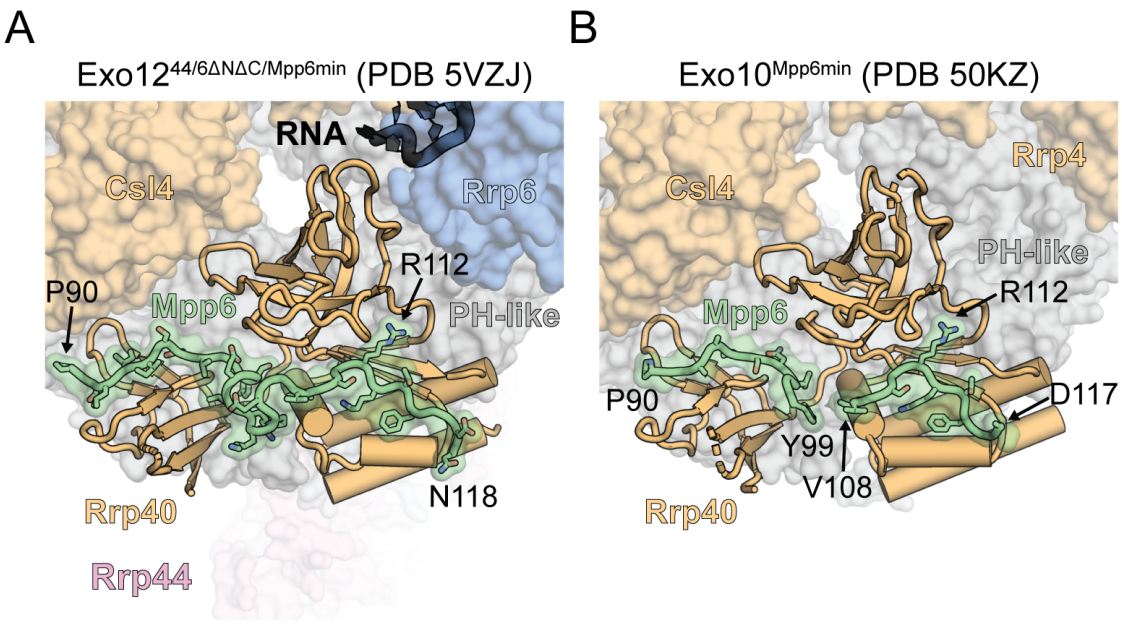


Figure 28. Mpp6$_{minimal}$ from two different crystal structures. Crystal structures of Mpp6 bound to the exosome core from this work (A) and Falk et al., 2017 (B). Terminal amino acids of the models and the R112 'arginine anchor' of Mpp6 are indicated. The structure in the asymmetric unit with the most residues modeled for Mpp6 from PDB 50KZ was aligned to Rrp40 from PDB 5VZJ.

Modeling of the W238R mutation in *S. cerevisiae* as well as mouse neurons revealed that the mutant protein was unstable and rapidly degraded by the proteaseome (Fasken et al., 2017). Human Exo9 reconstituted with Rrp40 W238R showed a decreased affinity for human Mpp6 in vitro (Falk et al., 2017), suggesting a loss of interaction between Mpp6 and the exosome may also contribute to the disease.

Mpp6 can recruit Mtr4 to the exosome

Results presented thus far do not explain the synthetic lethality observed when Mpp6 is deleted along with Rrp6, Rrp47, or Air1, a component of the TRAMP complex (Milligan et al., 2008). Mutagenesis experiments designed to disrupt the tripartite interaction between Rrp47-Rrp6 NTD-Mtr4^{Nterm} did not alter yeast viability, Mtr4-exosome association, or RNA processing (Schuch et al., 2014). However, synthetic lethality was observed when mutations in the tripartite interaction with Mtr4 were combined with a Mtr4 C-terminal GFP tag, consistent with Mtr4 recruitment to the exosome being dependent on additional factors or Mtr4 surfaces that interact with Rrp6, Mpp6, Rrp47 or other components of the TRAMP complex. Furthermore, experiments in human suggest that both Mpp6 and Rrp47 contribute to Mtr4 recruitment (Chen et al., 2001; Lubas et al., 2011; Schilders et al., 2007).

Mpp6 and/or Rrp47 were added to full-length Exo11$^{Dis3/Rrp6}$ to determine if these cofactors contribute to interactions with Mtr4 via analytical gel filtration (Figure 29). Consistent with previous results, Mtr4 did not co-elute with the exosome in the absence of Rrp47 and Mpp6 while co-elution was observed in the presence of Rrp47 (Schuch et al., 2014) (Figure 29). When Mpp6 and Rrp47 were both present, more Mtr4 was detected in fractions containing the exosome suggesting greater stability of the complex. Importantly, Mpp6 was capable of recruiting Mtr4 to the exosome in the absence of Rrp47 and, to lesser extent, in the absence of Rrp6. Removing the C-

terminal 30 amino acids of Mpp6 (1-156) did not disrupt Mtr4 interaction, but removal of the N-terminal 81 amino acids in Mpp6 (82-186) or the conserved N-terminal 22 amino acid motif in Mpp6 (23-120) resulted in no detectable interactions with Mtr4. $Mpp6^{Min}$ (81-120) failed to interact with Mtr4. These data suggest that Mpp6 elements required for exosome activation are distinct from those required for Mtr4 recruitment.

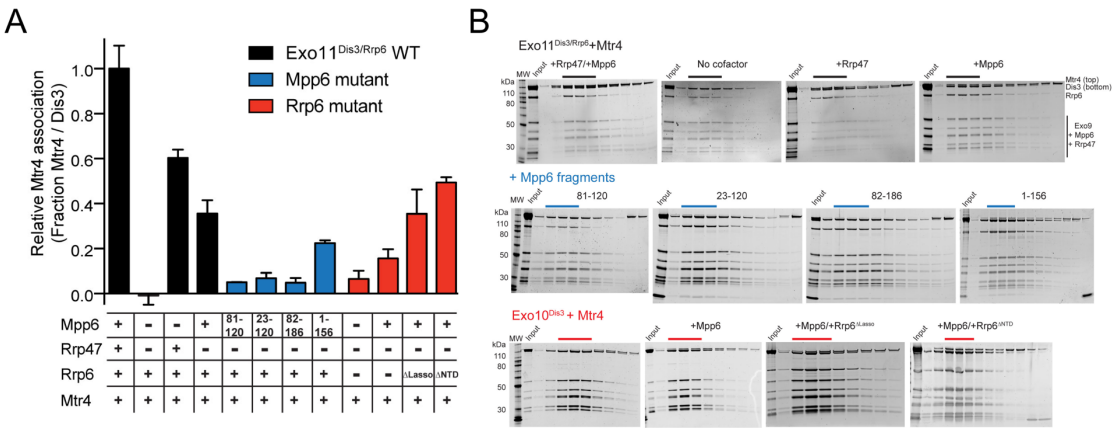


Figure 29. Mpp6 physically tethers Mtr4 to the nuclear exosome. (A) Mpp6 contributes to Mtr4 recruitment in Exo11$^{Dis3/Rrp6}$. Analytical gel filtration experiments were performed with Exo10^{Dis3} and Exo11$^{Dis3/Rrp6}$, Rrp47, Mtr4, with various truncations of Mpp6 and Rrp6 as indicated. Bar graphs represent the ratio of Mtr4 (122 kDa) to Dis3 (114 kDa) in peak fractions of the complex, as calculated by densitometric analysis of the fractions on SDS-PAGE, with error bars representing plus or minus one standard deviation. (B) Representative SDS-PAGE of analytical gel filtration runs (Superdex 200 Increase) of the indicated Mtr4-exosome combinations presented in panel A, with each lane representing an independent fraction. Gels were SYPRO Ruby stained and the ratio of Mtr4 to Dis3 was calculated via densitometry within lanes indicated by solid lines above each gel. Data acquired by Dr. Elizabeth Wasmuth with reagents prepared by John Zinder, Dr. Elizabeth Wasmuth, and Dr. Mom Das. Figure prepared by Dr. Elizabeth Wasmuth with assistance from Dr. Christopher Lima. Figure and caption adapted from Figure 5 and supplemental Figure 5 of Wasmuth et al., 2017.

Preparative reconstitution of the nuclear exosome using full-length components

Previous in vitro studies of exosome using the Rrp6/Rrp47 heterodimer have used C-terminally truncated of Rrp6 and Rrp47, which are particularly low-expressing and difficult to handle proteins, to facilitate their purification and subsequent characterization (Dedic et al., 2014; Makino et al., 2015; Schuch et al., 2014). We sought to develop methods for the purification of large quantities of highly pure full-length Rrp6/47 heterodimer, and while doing so re-optimized the purification of Rrp44, Mpp6, and Mtr4 proteins (Figure 30).

Rrp6 containing its N-terminal PMC2NT domain behaves poorly in the absence of its obligate binding partner (Dedic et al., 2014; Feigenbutz et al., 2013a), so both genes were cloned into a bacterial dual expression vector (see Chapter 1, Materials and Methods). An N-terminal His_6Smt3 tag was added to Rrp47 because free Rrp47 does not associate with the exosome (Stead et al., 2007), so excess protein can be easily separated from the complex in gel filtration after reconstitution (Figure 31A through 31C). Because the Rrp6 and Rrp47 C-termini bind nucleic acid (Costello et al., 2011; Wasmuth and Lima, 2017), we reasoned that using a heparin affinity column, the stationary phase of which is polyanionic and roughly resembles nucleic acid, after nickel elution would allow us to isolate full length proteins from C-terminal truncations. Integrity of the N-terminus would be assured via N-terminal tagging in the case of Rrp47 and the necessity of the N-terminus for interaction with Rrp47 in the case of Rrp6 (Stead et al., 2007). The His_6Smt3 tag was left on Rrp47 throughout its purification because we have empirically observed that it improves solubility. Adding Ulp during the exosome complex reconstitution allows for quantitative tag removal (Figures 31C and 31D). We were also able to successfully employ heparin affinity to purify His_6Smt3-Rrp44, His_6Smt3-Mtr4, and His_6Smt3-Mpp6 after nickel elution (see

Materials and Methods), though their Smt3 tags were cleaved prior to gel filtration (Figure 30).

Reconstitution of Exo13 was performed essentially as before, except one equivalent of Csl4, which is often substoichiometric in reconstituted complexes as well as those isolated from TAP-tagged yeast strains (Makino et al., 2013b; Wang et al., 2007), was added prior to dialysis to replace any that had been lost during the Exo9 reconstitution. Since Rrp6's C-terminus interacts extensively with Csl4, it stabilizes the protein on the complex (Makino et al., 2013b). Additionally, we used a 24 mL gel filtration column for reconstitution of 0.5-2 mg of complex rather than a 320 mL column as previously, which allowed for greatly increased time efficiency for FPLC use (40 min per run vs. 3 hrs per run) and for concentrating. For reconstitution of other complexes used in the next section, individual components were omitted from the reconstitution (e.g. for $Exo12^{Dis3/Rrp6/Rrp47}$ Mpp6 was omitted from the reconstitution). In complexes lacking Rrp47, Rrp6 contains a truncation of its N-terminal 128 amino acids to improve its stability (Feigenbutz et al., 2013a) and prevent its dimerization (Dedic et al., 2014).

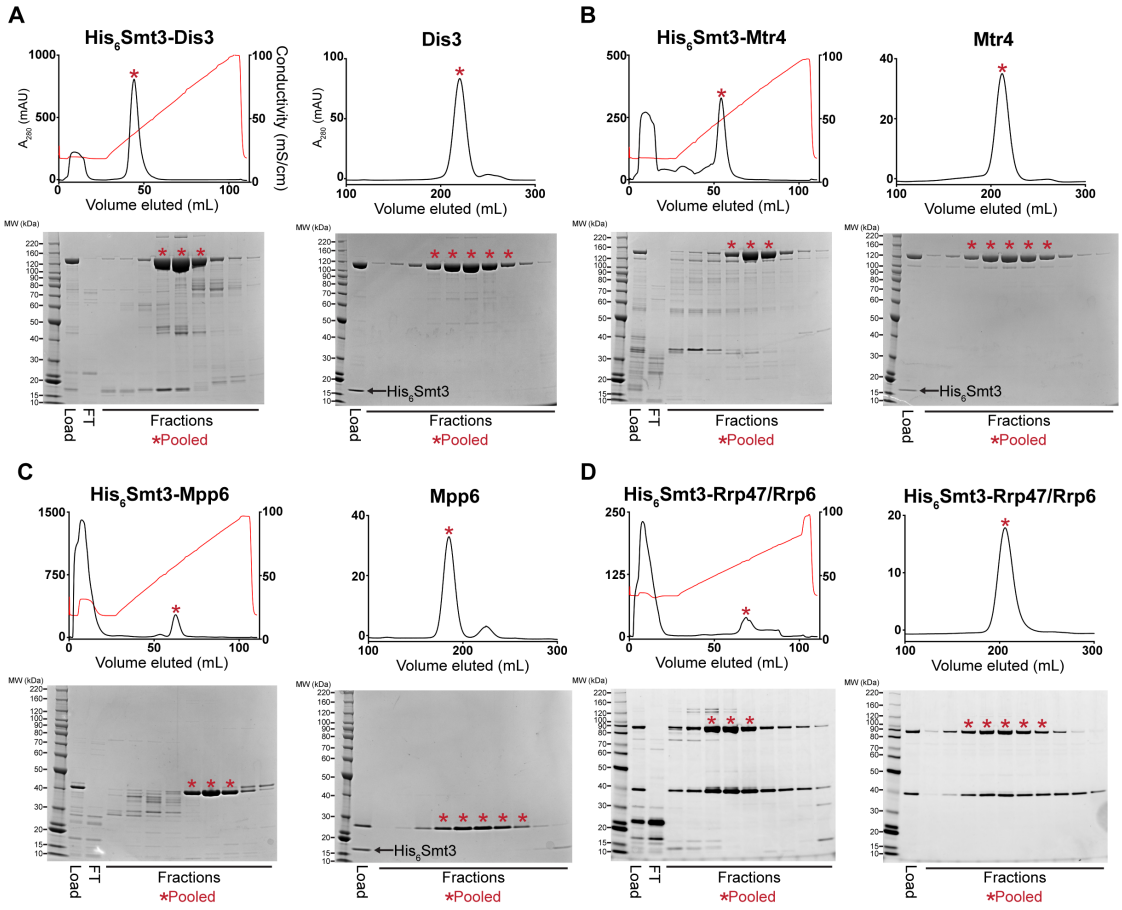


Figure 30. Purification of enzymatic subunits and Mpp6 cofactor of the nuclear exosome. Heparin affinity (left) and Superdex (right) purification of His_6Smt3 tagged (A) Dis3, (B) Mtr4, (C)Mpp6, and (D) Rrp47/Rrp6. The heparin purification was performed directly after elution from the nickel column, while the Superdex purification was run after overnight incubation of the heparin eluate with Ulp (A through C) or buffer (D). Traces show absorbance at 280 nm (A_{280}) (black curves) for heparin and Superdex columns and conductivity (red curves) for heparin only. Peaks of interested are indicated on the traces with an asterisk. SDS-PAGE analysis of the load, flowthrough (lanes labeled FT, for heparin only), and peak fractions are shown. Gels were stained with Coomassie (panels A through C) or SYPRO Ruby (panel D).

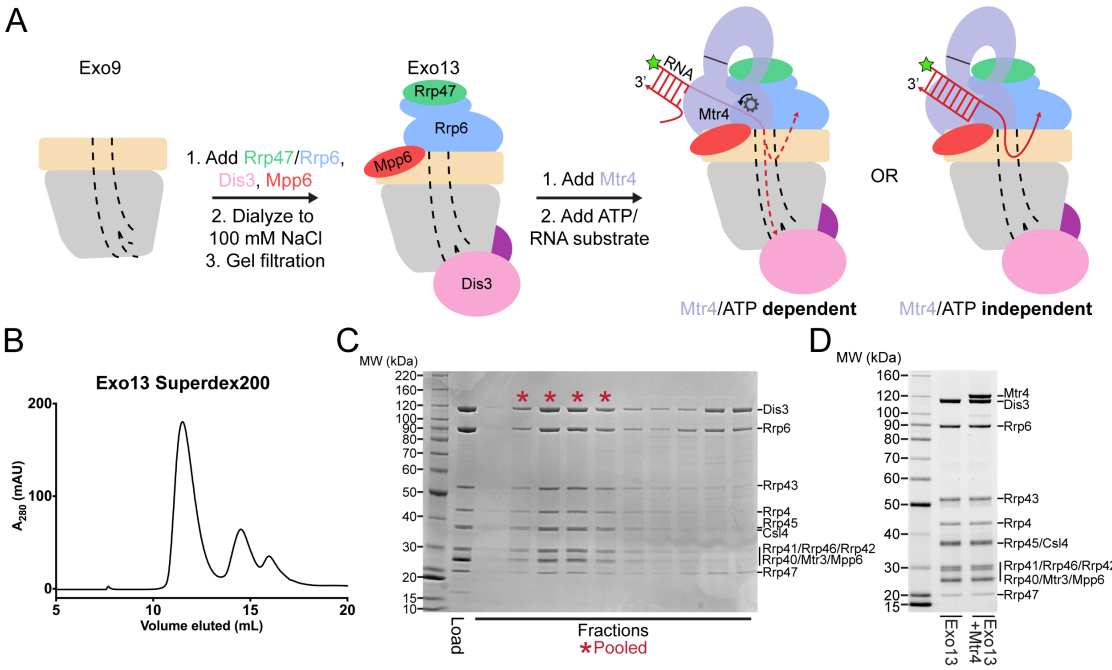


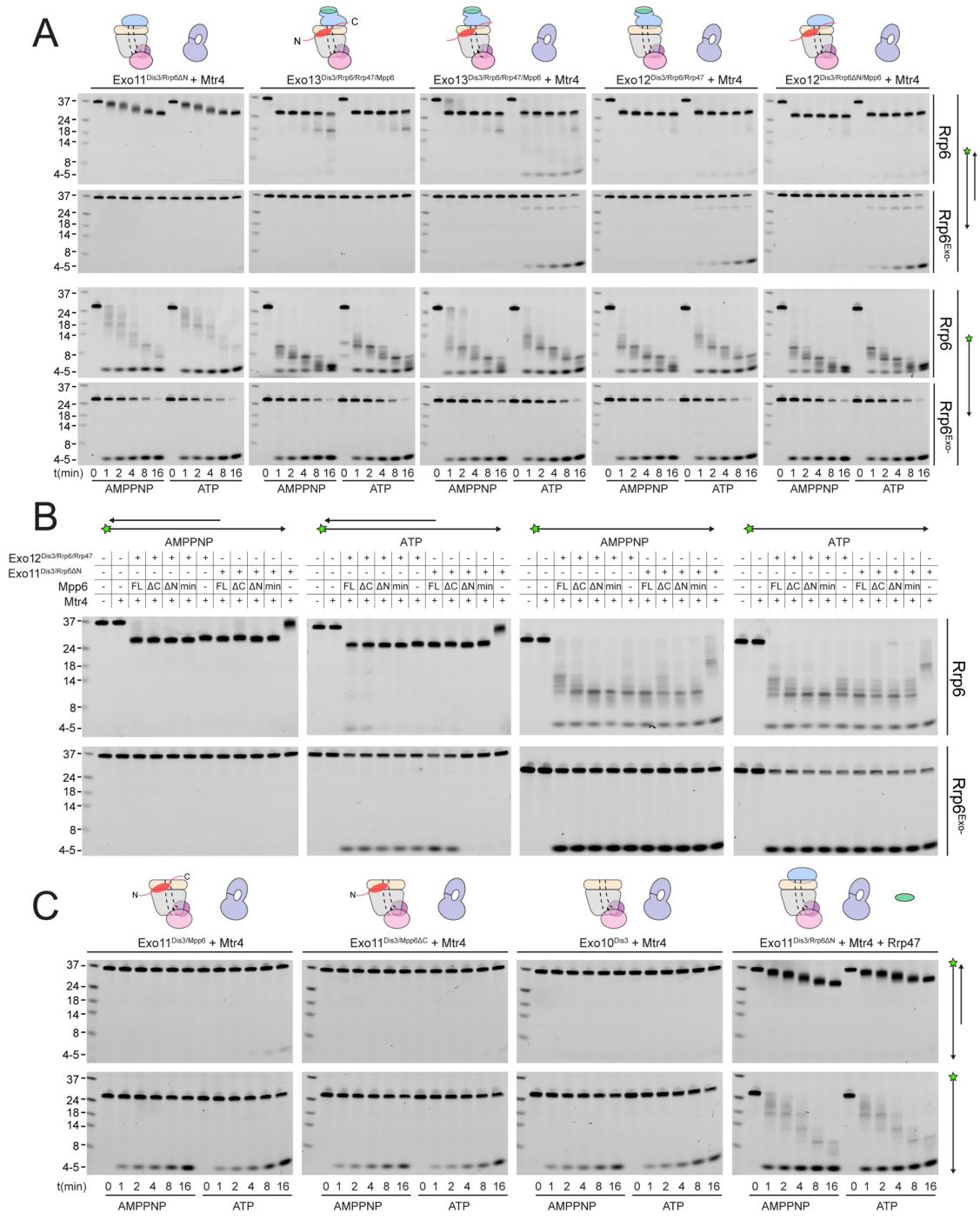
Figure 31. Strategy for reconstitution of Exo13 and of helicase-dependent RNA degradation. (A) Schematic depiction of Exo13 reconstitution strategy as well as Mtr4 mix-in for helicase dependent RNA decay. Shown are two possible RNA fates which are either dependent or independent of ATP and Mtr4 as indicated. The RNA substrate is shown as a red line with an arrow pointed towards the 3' end, and the 5' 6-carboxyfluorescein label is represented with a green star. Base pairs are represented as crossbars between parallel strands. (B) A_{280} trace and (C) Coomassie stained SDS-PAGE analysis of fractions from gel filtration of the Exo13 reconstitution. (D) SYRPO Ruby stained gel showing Exo13 with and without added Mtr4.

Reconstitution of Mtr4 and ATP dependent RNA decay

To assess whether Mpp6- or Rrp47-mediated recruitment of Mtr4 results in Mtr4-dependent RNA degradation, reconstituted exosomes were assayed for decay activity in the presence of exogenously added Mtr4 using a 17 nt double stranded RNA with a 10 nt 3' poly(A) overhang ($ds_{17}A_{10}$) (Makino et al., 2015; Schuch et al., 2014) in comparison with assays using the same labeled 27 nt RNA without the complementary 17 nt strand ($ss_{17}A_{10}$) (Figures 31A, 31D and 32). Previous studies reported that Dis3 is unable to degrade the $ds_{17}A_{10}$ substrate in the nuclear exosome (Makino et al., 2013b), presumably because the 3' overhang is too short to span the Exo9 central channel. We reasoned that after unwinding by Mtr4, the 3' end of the $ds_{17}A_{10}$ substrate may be able to reach Dis3 through the central channel.

Rrp6 trimming of the 3' overhang was stimulated in the presence of Rrp47, Mpp6 or both (Figure 32A), consistent with our biochemical observations using poly(A) RNA (Wasmuth et al., 2017). Products of Dis3 activity could be detected with the $ds_{17}A_{10}$ substrate, but only when Rrp47 and/or Mpp6 was present. Importantly, this activity was also dependent both Mtr4 and ATP, as AMP-PNP, a non-hydrolyzable ATP analog, failed to support activity. By contrast, Dis3 degradation of $ss_{17}A_{10}$ did not appear dependent on Mpp6, Rrp47, Mtr4, ATP or AMP-PNP. Trimming of the $ds_{17}A_{10}$ 3' overhang by Rrp6 dominated the assay in WT exosomes in the presence of Rrp47 or Mpp6, so we analyzed Mtr4-dependent decay using exosomes reconstituted with catalytically inactivated Rrp6 ($Rrp6^{exo-}$). As before, Mtr4-dependent decay was only observed in the presence of Rrp47 or Mpp6, while addition of both cofactors resulted in the greatest stimulation (Figure 32A). These results show that Mtr4 can promote degradation of dsRNA by the nuclear exosome, and that this activity is dependent on Mpp6 and Rrp47 cofactors.

Figure 32. Mtr4-dependent RNA degradation requires either Mpp6 or Rrp47. Urea-PAGE analysis of RNA decay products by indicated exosome complexes and cofactors. In reactions labeled 'Rrp6^{Exo-}', Rrp6 contains a D238N mutation to render its exonuclease site catalytically inactive. 5' fluorescein-labeled poly(A) RNA of indicated lengths are present in the leftmost lane of each gel with the Dis3 product labeled '4-5 nt'. Samples were not heated prior to gel electrophoresis, so the dsRNA substrate runs as a duplex at approximately 37 nt (A) Activity of reconstituted RNA exosome complexes in the presence or absence of equimolar exogenous Mtr4. (B) Mix-in of 1.5-fold molar excess of full-length Mpp6 or indicated truncations in RNA decay reactions containing reconstituted Exo12$^{Dis3/Rrp6/Rrp47}$ or Exo11$^{Dis3/Rrp6\Delta N}$ and equimolar Mtr4 (when present). Results shown for an end point assay after 1 minute incubation with exosomes containing Rrp6 or 8 minutes for exosomes containing Rrp6^{exo-}. (C) RNA degradation activities of reconstituted RNA exosome complexes and indicated mutations in the presence of equimolar exogenous Mtr4 and/or Rrp47. Cartoons shown above gels in panels A and C depict the exosome core in grey and wheat, Rrp6 in blue, Dis3 in pink, Mpp6 in red and Rrp47 in green. The central body of Mpp6 is shown as an ellipse with N- and C-terminal tails labeled or removed to reflect the protein used in the assay. The N-terminal PMC2NT domain of Rrp6 is shown as an appendage to the Rrp6 protein. From Figure 6 of Wasmuth et al., 2017.


A
Exo11Dis3/Rrp6ΔN + Mtr4
Exo13Dis3/Rrp6/Rrp47/Mpp6
Exo13Dis3/Rrp6/Rrp47/Mpp6 + Mtr4
Exo12Dis3/Rrp6/Rrp47 + Mtr4
Exo12Dis3/Rrp6ΔN/Mpp6 + Mtr4
Rrp6
Rrp6Exo-
t(min) 0 1 2 4 8 16 0 1 2 4 8 16
AMPPNP ATP
B
AMPPNP ATP
Exo12Dis3/Rrp6/Rrp47
Exo11Dis3/Rrp6ΔN
Mpp6 FL ΔC ΔN min
Mtr4
C
Exo11Dis3/Mpp6 + Mtr4
Exo11Dis3/Mpp6ΔC + Mtr4
Exo10Dis3 + Mtr4
Exo11Dis3/Rrp6ΔN + Mtr4 + Rrp47

Unwinding of the dsRNA region by Mtr4 appears responsible for Dis3 activity, but only when it is recruited to the exosome via Rrp47 or Mpp6. To determine the relevance of domains in Mpp6 required for this activity, we compared activities of exosomes in the presence of exogenous Mpp6 (FL: 1-186, ΔC: 1-120, ΔN: 81-186, and Min: 81-120) to those lacking Mpp6 in the presence of Mtr4 (Figure 32B). Results suggest that each Mpp6 construct stimulates Rrp6 trimming activity of the 3' overhang, but only FL and ΔC Mpp6 constructs, both of which recruit Mtr4 in gel filtration experiments (Figure 29), support ATP/Mtr4/Dis3-dependent RNA decay in the absence of Rrp47 (Figure 32B). In the presence of Rrp47, Mpp6 is largely dispensable for this activity, but FL and ΔC Mpp6 further increased Mtr4-dependent decay. As before, similar trends were observed in assays conducted with exosomes containing $Rrp6^{exo-}$.

Perhaps consistent with the observation that yeast strains are viable in the absence of Rrp6 and Rrp47 (Feigenbutz et al., 2013a; Stuparevic et al., 2013), some Mtr4-dependent decay could be observed in complexes lacking Rrp6 (Figure 32C), although this required full-length Mpp6 and its activities were weaker in comparison to exosomes containing Rrp6 or Rrp6 plus Rrp47. Furthermore, addition of Mpp6 to exosomes lacking Rrp6 or Rrp47 also showed an ability to interact with Mtr4 (Figure 29). These data suggest that Mpp6 could contribute to Mtr4 recruitment in the absence of Rrp6 and Rrp47 proteins (Schuch et al., 2014). Importantly, ATP/Mtr4/Dis3-dependent activity on the $ds_{17}A_{10}$ substrate was only observed in complexes that could physically interact with Mtr4 (Figures 29 and 32), suggesting that the unwinding activities of Mtr4 lead to productive degradation only when Mtr4 is physically tethered to the RNA exosome.

Optimal cell growth depends on unique domains in Rrp6 and Mpp6

Genetic interactions between Mpp6, Rrp47 and Rrp6 have been reported previously (Feigenbutz et al., 2013b; Milligan et al., 2008). While none of these genes is essential, deletion of Mpp6 and Rrp6 results in synthetic lethality. Furthermore, viability is maintained in cells lacking both Rrp47 and Rrp6 as long as Mpp6 remains present. To relate the relevance of Mpp6 and Rrp6 domains defined in this study as important for exosome stimulation and Mtr4 recruitment to function in vivo, strains lacking Rrp6 and Mpp6 were complemented with mutant alleles for Rrp6 and/or Mpp6 in various combinations (Figure 33). Western blotting of these strains showed expression of Rrp6 in all cases, with especially high expression for N-terminal truncations as previously observed (Figure 34) (Wasmuth and Lima, 2017). We also attempted to raise polyclonal antibodies against $Mpp6_{minimal}$ but failed to achieve sufficient signal-to-noise to detect even the WT protein in whole cell lysates, possibly because of His_6Smt3 tag contamination in the antigen sample.

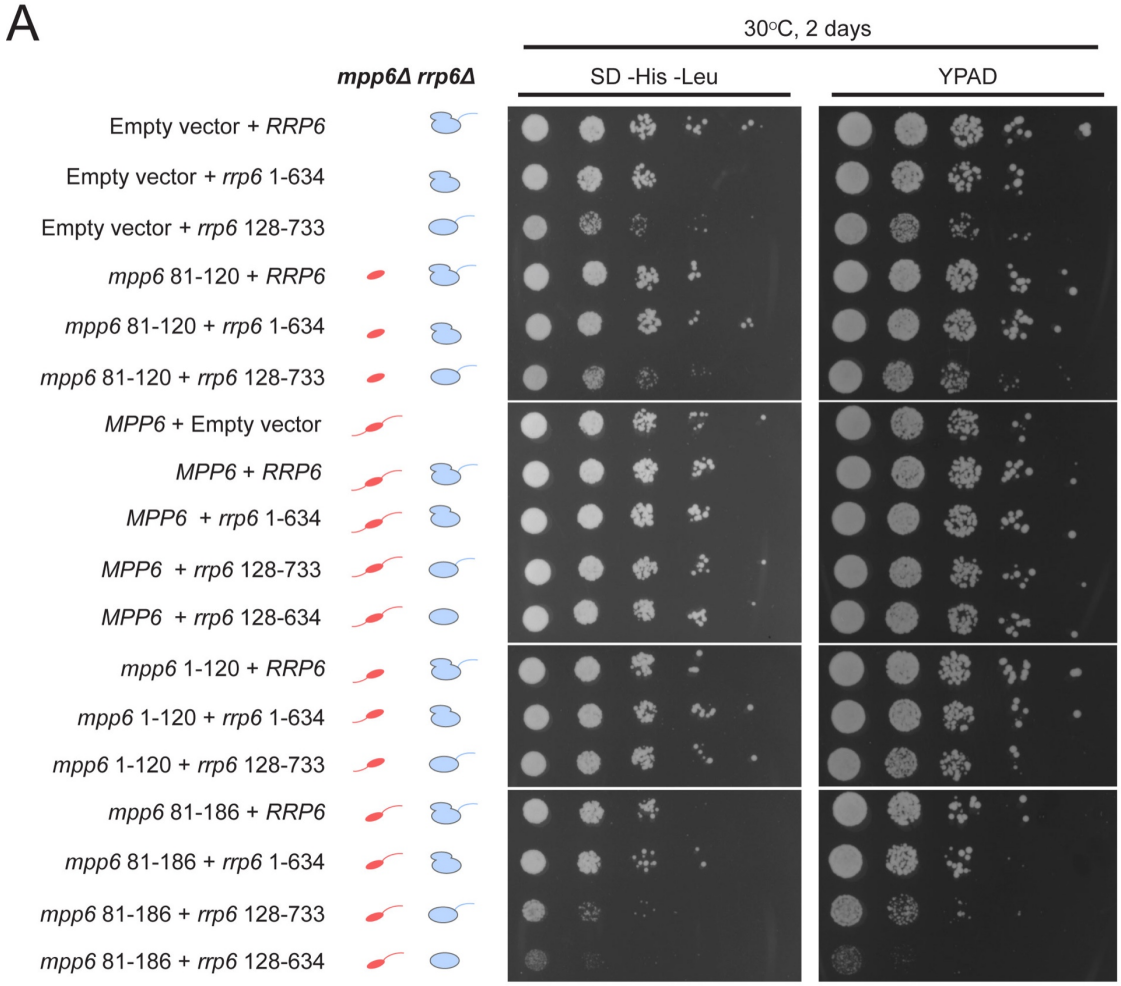
As stated previously, strains lacking both Rrp6 and Mpp6 are not viable while strains containing either full-length Mpp6 or full-length Rrp6 are viable. Importantly, in strains lacking Mpp6 removal of the PMC2NT domain in *rrp6 (128-733)*, a combination that should eliminate Rrp6 interactions with Rrp47, and by extension the Mtr4 helicase, resulted in slower growth while removal of both the Rrp6 PMC2NT and lasso in *rrp6 (128-634)* resulted in no viable colonies (Figure 33B). In contrast, in strains lacking Rrp6, only full-length Mpp6 suppressed lethality as strains complemented with Mpp6 truncations *mpp6 (1-120)*, *mpp6 (81-186)* and *mpp6 (81-120)* were not viable (Figure 33). This observation is consistent with our in vitro data, namely that in the absence of Rrp6, full-length Mpp6 can still mediate recruitment of Mtr4 to the exosome (Figures 29 and 32C). In addition, growth defects associated with *rrp6 (128-733)*, an allele that should not interact with Rrp47, are suppressed by WT

Mpp6 and *mpp6 (1-120)*, but not other *mpp6* variants. This result is perhaps consistent with biochemical observations, namely that Mpp6 residues 1-120 contribute to recruitment of Mtr4 and Mtr4-dependent decay in the absence of Rrp47 (Figures 29 and 32B).

Strains containing *mpp6 (81-120)* phenocopied Δ*mpp6*, suggesting that either this construct is not expressed or that the exosome association domain is not sufficient to support Mpp6 functions in vivo, despite its ability to stimulate nuclear exosome activity in vitro (Wasmuth et al., 2017). However, since the Mpp6 (81-120) protein did not result in productive interactions with Mtr4 in vitro (Figures 29), the observation that *mpp6 (81-120)* phenocopies Δ*mpp6* is also consistent with a model in which Mtr4 recruitment is defective. Finally, a strain harboring *mpp6 (81-186)* and *rrp6 (128-634)* was viable despite the fact that both Rrp47- and Mpp6-dependent mechanisms for Mtr4 recruitment are presumed missing. Although this strain grows much slower than any other characterized in this study, its viability suggests that combining the Mpp6 exosome associating region with its C-terminal domain is sufficient to suppress lethality observed for *rrp6 (128-634)*, perhaps through interactions with the transcriptional termination machinery via Nrd1 as noted previously (Kim et al., 2016).

Figure 33. Optimal cell growth depends on unique domains in Rrp6 and Mpp6. (A) Growth analysis of *S. cerevisiae* strains carrying viable combinations of Rrp6 and Mpp6 alleles. Ten-fold dilutions series of the *rrp6Δ mpp6Δ* strains transformed with the indicated pRS415 *mpp6* and pRS413 *rrp6* plasmids. Spotting was performed on SD-His-Leu or YPAD solid media and cells were incubated at 30°C for 2 days. Cartoons depict *mpp6* and *rrp6* alleles with respect to N- and C-terminal deletions. (B) Scoring table of the yeast growth phenotypes established in panel A after 1 day growth. Scoring is based on a subjective five-point system where five "+" symbols ("+++++") correspond to the fastest observed growth rate and one "+" symbol to the slowest growth rate. Plasmid combinations indicated with the "-" symbol resulted in synthetic lethality. Data generated by Dr. Dimitrios Zattas. Figure prepared by Dr. Dimitrios Zattas with assistance from Dr. Christopher Lima. Figure 7 of Wasmuth et al., 2017.

A



B

	pRS413 (empty)	*RRP6*	*rrp6* 1-634	*rrp6* 128-733	*rrp6* 128-634
pRS415 (empty)	-	++++	+++	++	-
MPP6	+++	+++++	+++++	+++	+++
mpp6 1-120	-	++++	+++	+++	-
mpp6 81-186	-	++++	+++	++	+
mpp6 81-120	-	++++	+++	++	-

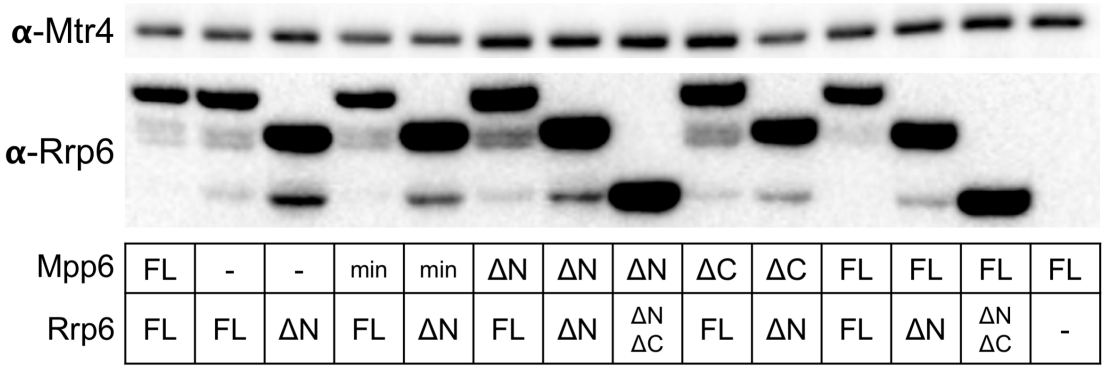


Figure 34. Expression of Rrp6 variants. Western blot of soluble *S. cerevisiae* lysates after cryo-milling and extraction using affinity purified rabbit polyclonal antibodies against *S. cerevisiae* Rrp6 residues 128-518 or *S. cerevisiae* Mtr4 residues 80-1073. From selected strains in Figure 33. After primary antibody treatment, membranes were incubated with anti-rabbit-horseradish peroxidase, washed, developed, and visualized for chemiluminescence. For Mpp6, FL encompasses residues 1-186, min is 81-120, ΔN is 120-186, and ΔC is 1-120. For Rrp6, FL encompasses residues 1-733, ΔN is 128-733, and ΔNΔC is 128-634.

Discussion

Experiments presented here suggest that the nuclear exosome cofactor, Mpp6, can associate with the Exo9 core through interactions with the S1/KH subunit Rrp40 to recruit the helicase Mtr4 and to stimulate Rrp6 activity in the exosome. A minimal region of Mpp6 comprised of amino acids 81-120 appears sufficient for exosome association and stimulation in vitro, although it is defective for recruitment of Mtr4. As our structure shows, this fragment includes a conserved arginine anchor that is important for Mpp6 interactions with the exosome core.

This work also sheds light on dual functions of Mpp6 and Rrp47 insofar as both factors can stimulate RNA exosome activities and both contribute to Mtr4 recruitment and Mtr4-dependent degradation in vitro, perhaps providing some additional insights to in vivo synthetic interactions reported in previous studies (Feigenbutz et al., 2013b; Milligan et al., 2008). Interactions between the nuclear exosome and Rrp47 or Mpp6 result in stimulation of exosome activities, but they rely on distinct surfaces for these activities. Rrp47 interacts with the N-terminal PMC2NT domain of Rrp6 (Makino et al., 2015; Schuch et al., 2014; Stead et al., 2007), while Mpp6 interaction occurs via the Exo9 core (Falk et al., 2017; Schuch et al., 2014).

Rrp47 and Mpp6 each contribute to Mtr4 interactions with $Exo11^{Dis3/Rrp6}$, and inclusion of both cofactors appears to maximize these interactions. While we were unable to decouple Rrp47-mediated stimulation from its ability to recruit Mtr4, genetic and biochemical results suggest that a minimal fragment of Mpp6 (residues 81-120) that supports Mpp6-exosome interaction was not sufficient to support growth in vivo when paired with an *rrp6* allele that does not interact with Rrp47, and that additional Mpp6 elements are required for optimal cell growth.

Results from our characterization of cofactor-dependent Mtr4 recruitment for degradation of a structured substrate show that Rrp6's activity dominates the reaction and creates a trimmed product that is apparently stable towards further degradation, even in the presence of Mtr4 and ATP (Figure 32A). In contrast, the substrate is nearly completely degraded by Dis3 over the course of the reaction when a catalytically inactive Rrp6 is present using the same conditions (Figure 32A), suggesting that Rrp6's activity can serve to protect structured substrates from full degradation by Dis3 in some cases. In cells, Mtr4 can be incorporated into the TRAMP complex (LaCava et al., 2005; Wyers et al., 2005), the polyadenylation activity of which could antagonize this protective trimming by Rrp6 and allow Mtr4/Dis3 additional opportunities to engage the 3' end. Furthermore, polyadenylation by TRAMP is stimulated on unmodified or mutant substrates (Vanacova et al., 2005), providing the model for selective degradation of these substrates over their healthy counterparts by the TRAMP-exosome presented in Figure 35. One caveat of this simplified model is that it does not hold for all RNAs or in all organisms, especially in light of findings that human Rrp6's activity has been implicated in the total degradation of some substrates in human and that Dis3's activity can be used for processing and not degradation in both yeast and human (Fromm et al., 2017; Tomecki et al., 2014) (Figure 8D). Additionally, the poly(A) polymerase and zinc-knuckle components of TRAMP are found primarily in the nucleolus of human, where Dis3 is depleted (Lubas et al., 2011; Staals et al., 2010; Tomecki et al., 2010) (Table 1 and Figure 1B), suggesting different means for RNA quality in higher eukaryotes.

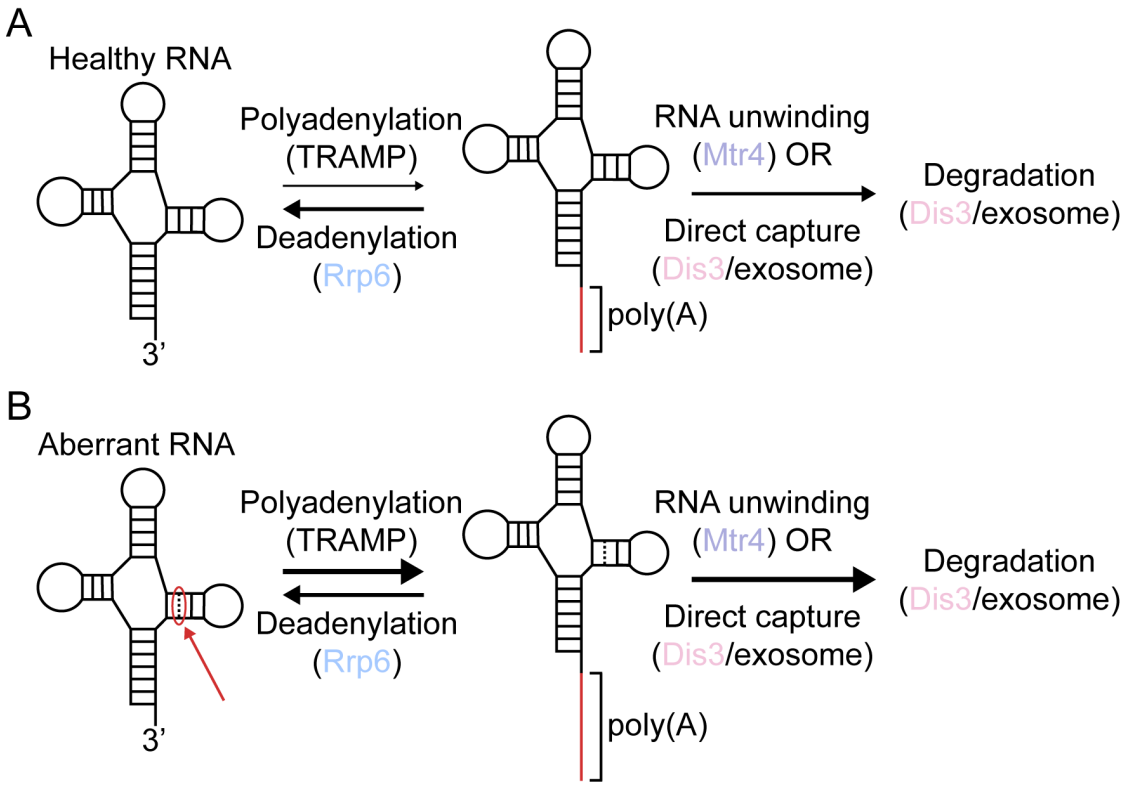


Figure 35. Model for RNA quality control by the *S. cerevisiae* nuclear RNA exosome. A prototypical structured RNA is represented as a black line and base pairs between strands shown as crossbars. Arrow thickness indicates the predicted favorability of a given transition. 'Direct capture' refers to the possible scenario in which polyadenylation alone can generate a poly(A) tail long enough to engage Dis3 without need for unwinding by Mtr4. (A) Healthy RNAs are deadenylated and unwound at lower rates than aberrant RNAs, so flux through those pathways is limited. (B) Aberrant RNAs will be polyadenylated and unwound at faster rates, stimulating their degradation. A generic aberration (e.g. lesion, mutation, lack of post-transcriptional modification) is indicated by a dashed line and highlighted with a red arrow.

Our genetic experiments uncovered the curious finding that some strains that include deletions and truncations that cannot recruit Mtr4 in vitro, for example *rrp6(128-733)*Δ*mpp6*, showed only a modest growth defect and not lethality as in Δ*mpp6*Δ*rrp47*. One possible explanation is that recruitment of Mtr4 to the exosome is not essential for viability. If this is the case, Rrp47 may be playing roles outside of the context of the exosome that are essential in the context of Mpp6 knockout. This seems unlikely, however, as Rrp47 is rapidly degraded in yeast strains that lack Rrp6's N-terminus (Feigenbutz et al., 2013b). Another possibility is that increased expression of Rrp6 when its N-terminus has been removed can compensate for poor Mtr4 recruitment in vivo (Wasmuth and Lima, 2017) (Figure 34). If this is the case, Rrp47's essential functions in *mpp6* null strains could be to stabilize Rrp6 from degradation (Feigenbutz et al., 2013a), to recruit Mtr4 (Schuch et al., 2014), and/or to stimulate Rrp6's activity. A third and related explanation is that there could be other yet unknown factors that can facilitate helicase-dependent RNA decay by the exosome in the absence of Rrp47 and Mpp6 that require stable Rrp6 bound to the core, which would not be the case if Rrp6 contained its N-terminus in the absence of Rrp47 (Feigenbutz et al., 2013a). Additional details have been resolved by our work with respect to the contributions of Mpp6 and Rrp47 to nuclear exosome function, however critical work remains to resolve the structural and mechanistic bases for activities associated with the Mtr4 helicase.

Materials and Methods

Exosome subunit purification, complex reconstitution, and analyses.

Cloning, expression, purification of exosome subunits and reconstitution of various complexes have been described previously (see Chapter 1, Materials and Methods). *S. cerevisiae* Mpp6 (full-length and various truncations), Rrp47, and Mtr4 were cloned into pRSF-Duet1 with a N-terminal Smt3-fusion tag, and were transformed into *E. coli* BL21 (DE3) RIL (Novagen). Recombinant protein expression was induced by addition of 0.4 mM isopropyl-β-D-thiogalactoside and overnight shaking at 18°C. Cells were lysed by sonication, and supernatants were purified by Ni-NTA. For Mpp6 variants and Rrp47, further purification was performed on a Superdex 75 (GE), followed by overnight cleavage of the Smt3 tag using Ulp1 protease, and a final purification on a heparin Hi-Trap column to remove Smt3 and nucleic acid impurities. Mtr4 was purified on a Superdex 200 (GE), subject to Ulp1 cleavage overnight, and purified by MonoQ (GE) to remove Smt3. Alternative purification protocols for Dis3, Mpp6, and Mtr4 are described in Figure 30 were employed for the reconstitution of complexes used for RNA decay (Figure 32). Briefly, these proteins were eluted from the nickel resin using a low-salt buffer (20 mM Tris-Cl, 150 mM NaCl, 250 mM imidazole-HCl pH 8.0, 1 mM BME) and injected over heparin column. The column was washed and the proteins eluted in using a linear gradiant from 150 mM to 1 M NaCl in 20 mM Tris-Cl pH 8.0 1 mM BME. Proteins were then Ulp digested overnight at 4°C and purified over Superdex200 (for Rrp44 and Mtr4) or Superdex75 (for Mpp6) to remove the cleaved tag.

For interactions with Mtr4 (Figure 29), 1.0 μM $Exo11^{Dis3/Rrp6}$ was incubated with 5-fold molar excess Mpp6 or Rrp47, and 2-fold molar excess Mtr4 on ice for 30 minutes in a total volume of 400 microliters of 100 mM NaCl, 40 mM MES pH6.5, 1

mM TCEP and run on a Superdex 200 Increase (GE), and fractions analyzed by SDS-PAGE and stained with Sypro Ruby. Densitometry was used to calculate the ratios of Mtr4 to Dis3 in fractions corresponding to nuclear exosome peaks using Fujifilm Multi Gauge.

3'-3' RNA synthesis and purification.

Alkynyl (5' hexynyl UUU AUU AUU UAU UUU AAA A 3') and azido (5' azide/NHS UUA UUU UAA AA 3') RNAs were synthesized and purified by Integrated DNA Technologies. RNAs were ligated by incubating 200 μM alkynyl RNA with 100 μM azido RNA in 80 mM potassium phosphate (pH 7.0), 5 mM sodium ascorbate, 0.5 mM $CuSO_4$/THPTA at 25°C for 1 hour. Quenching and purification of the reaction by DEAE chromatography were performed as previously described (see Chapter 1, Materials and Methods).

Crystallization and structure determination.

$Exo12^{Dis3exo\text{-}endo\text{-}/Rrp6exo\text{-}/Mpp6}$ was mixed in a 1:1 molar ratio with a single-stranded 30 nucleotide RNA bearing two 3' ends synthesized by click chemistry (See Chapter 1) incubated on ice for 30 minutes. After one week, the protein-RNA complex formed single crystals at 4°C in 11-13% PEG3350, 100 mM NaCitrate pH 5.6, 7 mM MES pH 6.5, and 175-200 mM ammonium sulfate that continued to grow for up to one month. Crystals were cryoprotected in mother liquor augmented with 21% glycerol, introduced in three steps of 7% increments. Native x-ray diffraction data were collected at the Advanced Photon Source 24-ID-E and 24-ID-C beam lines (NECAT), and 23-ID-D (GM/CA) at the selenium edge. Data were processed using HKL2000 (Otwinowski and Minor, 1997) and the structure solved by molecular replacement using Phaser (McCoy et al., 2007) in the Phenix suite (Adams et al.,

2010) with the coordinates of yeast Exo10$^{Rrp6exo-}$ (PDB: 4OO1) and Exo11$^{Dis3exo-endo-/Rrp6exo-}$ (PDB: 5K36) as search models (Table 2). Iterative rounds of refinement were accomplished using Phenix (Adams et al., 2010). RNA and protein were manually built using Coot (Emsley et al., 2010). The model was refined using positional refinement, real-space refinement and individual B-factor refinement. Figures depicting structures were prepared with PyMol (Schrödinger). Surface conservation was calculated using ConSurf (Ashkenazy et al., 2010). Structure quality was assessed using MolProbity (Chen et al., 2010).

RNA degradation assays.

RNA oligonucleotides were synthesized with a 5' fluorescein as described previously (Liu et al., 2006; Wasmuth et al., 2014) and purchased from Integrated DNA Technologies (IDT) or Dharmacon. For the substrates used in Figure 32, the 5' 6-carboxyfluorescein bottom strand (5' FAM CCC CAC CAC CAU CAC UUA AAA AAA AAA 3') and duplex top strand (5' AAG UGA UGG UGG UGG GG 3') were synthesized and HPLC purified by IDT. To prepare the $ds_{17}A_{10}$ substrate used in Figure 32, top and bottom strands were mixed at 100 and 150 µM, respectively, in 20 µL of 10 mM Tris-Cl pH 8.0, 100 mM KCl, 0.5 mM EDTA, heated to 95°C for 5 min, cooled to 60°C for 2 minutes, then to 4°C in a thermocycler. The $ss_{17}A_{10}$ was prepared in a similar manner, except the top strand was excluded from the reaction. Annealing reactions were fractionated by DEAE chromatography (Waters Protein-Pak DEAE 8HR 1000 Å 8 µm 5x50 mm column) using a linear gradient from 300 mM NaCl to 700 mM NaCl in 10 mM Tris-Cl pH 8.0, 0.1 mM EDTA pH 8.0 at 40°C. Fractions were collected and analyzed by TBE-PAGE using SYBR Gold staining (Life Technologies). RNA in fractions containing the desired product was ethanol precipitated and the pellet resuspended in ultrapure H_2O.

For the RNA degradation assays in Figure 32, the final concentrations were 10 nM exosome, 10 nM Mtr4 if present, 1 mM ATP or AMPPNP, 0.5 U/µL RNAse inhibitor (New England Biolabs), and 10 nM RNA substrate in 20 mM HEPES-KOH pH 7.5, 50 mM KCl, 1.1 mM $MgCl_2$, 0.5 mM TCEP-KOH pH 7.0, 0.02 % IGEPAL co-630. 50 mM ATP and AMPPNP stock solutions were adjusted to pH 7.0 with KOH. Exosome complexes were incubated on ice with Mtr4 at 1 µM each in the above buffer for 1 hour prior to initiating the reaction. A mix containing RNA, ATP or AMPPNP, and RNAse inhibitor (all at 1.1x final concentration) was incubated at 20°C for 5 min prior to initiation with 1/10 volume of 100 nM enzyme. Reactions were quenched after the indicated incubation times by adding 10 µL of reaction to 5 µL of stop mix (0.3 % w/v SDS, 30 mM EDTA pH 8.0, 3 U/mL proteinase K [New England Biolabs]) followed by proteinase K digestion at 37°C for 1 hour and flash freezing in liquid nitrogen for storage at -80°C. Prior to gel loading, 15 µL of 89 mM Tris-borate pH 8.3, 7 M Urea, 2 mM EDTA, 12% w/v Ficoll, 0.005 % w/v xylene cyanol was added. The Mpp6 mix-in experiments were carried out in a similar fashion, except the indicated Mpp6 construct was included in the initial enzyme mix at 1.5 µM (1.5-fold molar excess). These reactions were incubated at 20°C for 1 or 8 min for Rrp6 and $Rrp6^{exo-}$, respectively. 7.5 µL was loaded onto 15% acrylamide TBE urea gels (Life Technologies) to detect decay products. Gels were imaged for fluorescein fluorescence using a Typhoon FLA9500 instrument (GE).

Yeast growth assays.

Detailed protocols for generation of the strains can be found in (Wasmuth et al., 2017). Briefly, to obtain viable *rrp6 mpp6Δ* strains in the presence of various combinations of pRS413 *rrp6* and pRS415 *mpp6* vectors, the *rrp6 mpp6Δ* + pRS426 *RRP6* strain was transformed with these plasmids and plated on selective SD -Ura-

His-Leu solid medium. Individual clones were then streaked onto SD-His-Leu+5-FOA medium to select for viable strains lacking the pRS426 *RRP6* cover plasmid but maintaining the pRS415 and pRS413 vectors. Growth was monitored at 25°C and 30°C for 8 days. Yeast growth via serial dilution was performed based on a recently published protocol (Watts et al., 2015). All yeast strains were grown in selective SD -His-Leu medium overnight at 30°C. The next day, cells were diluted to an OD_{600} of 1 before being spotted on SD -His-Leu or YPAD media in a ten-fold dilution series using a multichannel pipette (4 µl per spot). Cells were incubated at 30°C for 2 days and then imaged.

Antibodies and Western Blotting

Polyclonal antibodies were generated by inoculating rabbits with either purified *S. cerevisiae* Rrp6 residues 129-518 (Wasmuth and Lima, 2017) or Mtr4 residues 80-1073 (Pocono Rabbit Farm). Affinity purifying the resulting sera generated primary antibody stocks that were used at 1:2500 dilutions. Western blots were developed by incubation with a 1:5000 dilution of anti-rabbit IgG-HRP (GE) followed by washing and treatment with ECL solutions (Pierce). Membranes were imaged using a GelDoc (Bio-Rad).

Cell powders of the indicated strains were prepared from 2L of culture grown in SD -His-Leu at 30°C with shaking to an OD_{600} of 1.0 by grinding in a Cryomill (Retsch) at liquid nitrogen temperatures. Powder was stored at -80°C until needed. For protein extraction, cell powder was resuspended in 3 volumes of buffer (50 mM HEPES-NaOH pH 7.5, 100 mM NaCl, 0.5 mM $MgCl_2$, 0.1% v/v IGEPAL ca-630, 5 mM TCEP-NaOH pH 7.0, 1x EDTA free protease inhibitor from Pierce, 1 mM PMSF) by briefly vortexing followed by incubation on ice for 10 minutes. The suspension was then centrifuged at 20,000 x g for 10 minutes and 10 µL of clarified lysate was added

to 65 µL of ddH_2O and 25 µL of 4x LDS loading buffer containing 100 mM BME. These samples were boiled for 5 minutes, then 5 µL was fractionated by SDS-PAGE, transferred to a PVDF membrane, and treated with antibodies using standard protocols.

CHAPTER 3: Structural Studies of Helicase-Dependent RNA decay by the Mtr4-Exosome Complex

Introduction

RNA helicases play central roles in many areas of RNA biology including RNA processing, RNA decay, and RNP biogenesis (Pyle, 2008). The related helicases Ski2 and Mtr4 both contribute to RNA exosome function as members of the Ski and TRAMP complexes, respectively (Figures 1, 6, and 7). In *S. cerevisiae*, essentiality of Mtr4's ATPase activity and synthetic lethality of Δ*ski2*Δ*xrn1* emphasize the importance of these helicases in the RNA exosome's functions (Johnson and Kolodner, 1995; Taylor et al., 2014). Estimates of protein copy number using complementary methods place *S cerevisiae* Mtr4 at higher protein expression levels than the other components of the TRAMP polyadenylation complex (Ghaemmaghami et al., 2003; Kulak et al., 2014). This combined with the observation that disrupting Mtr4/Trf interactions is non-lethal (Losh et al., 2015) suggests that Mtr4 can act either inside or outside of the context of TRAMP.

Mtr4 is recruited to the *S. cerevisiae* exosome through the Mpp6 cofactor and the Rrp6/Rrp47 heterodimer (Schuch et al., 2014; Wasmuth et al., 2017) (Figure 29), and this recruitment enables degradation of structured substrates (Figure 32). Crystal structures of Mtr4 and exosome complexes have lead to a speculative structural model for channeling by the TRAMP-exosome complex presented in Figures 6B and 6C, however the connections that tether Mtr4 to the exosome as well as how the orientation of the helicase relative to the exosome remain largely unknown. To address this, we sought to structurally characterize a 14-component exosome complex (Exo14) from *S. cerevisiae* containing the Exo9 core, Rrp6, Rrp47, Dis3, Mtr4 and an

engineered substrate. While the work is ongoing, we observe a 3D class at an overall resolution of 4.0 Å into which exosome components and archless Mtr4 can be readily docked. In the model, Mtr4 sits atop the S1/KH cap using the same surface as Rrp6 does in other structures (Makino et al., 2015; Wasmuth et al., 2014; Zinder et al., 2016), and is channeling RNA through Exo9 to Dis3. Some density corresponding to the Rrp6 EAR (Figure 4A) is observed, but the majority of Rrp6/Rrp47 as well as Mpp6 cannot be docked into the volume. We also observe a less populated class at approximately 6.0 Å which contains Mtr4's arch and additional densities that may correspond to Rrp6's N-terminal domains and possibly Rrp47. Finally, we provide evidence that exosome complexes capable of recruitment of Mtr4 can stimulate its helicase activities for degradation of structured substrates.

Results

Design and testing a substrate for capturing the Mtr4-exosome

Our ultimate goal was to trap the Mtr4-exosome in a state after unwinding of the substrate and channeling to Dis3 to degradation. One of the challenges we faced was that an RNA 3' end has multiple components to engage in this complex, namely Mtr4, Dis3 (if the ssRNA portion is sufficiently long) and Rrp6. To overcome this, we include in our substrate a double stranded region with a short (8 nt) poly(A) overhang that we predicted would predominantly bind Mtr4 when incubated with the complex in the absence of ATP (Figure 36).

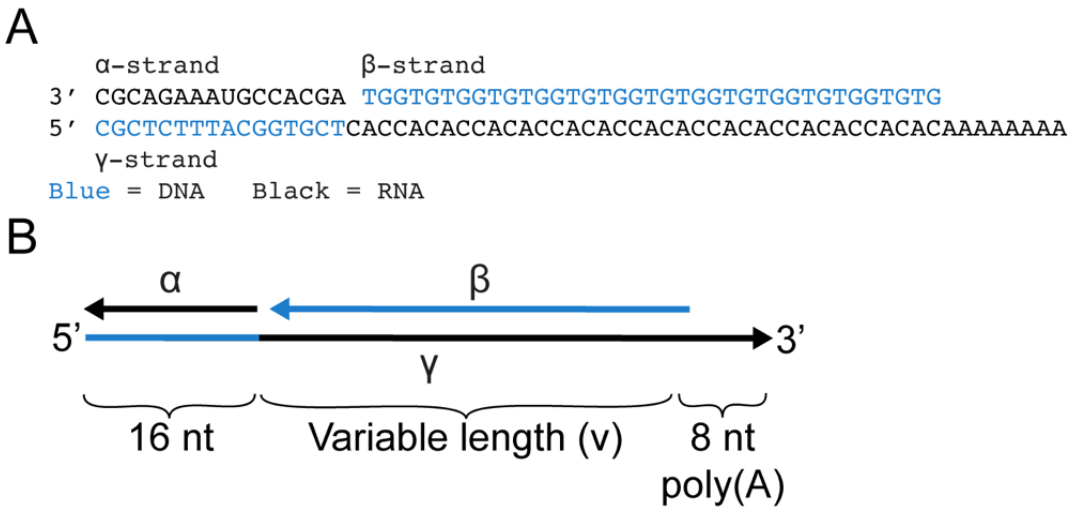


Figure 36. Substrate for capturing loaded Exo14. (A) Example nucleotide sequence for a tripartite substrate with length of the variable region (v) equal to 38 nucleotides. (B) Schematic representation of tripartite substrates. Black represents RNA regions, blue represents DNA, and arrows point towards the 3’ end.

Another challenge we faced was that we wished to stall the helicase after ATP had been added to prevent dissociation. Mtr4 is a DExH family helicase that can use ATP to translocate along the 'bottom strand,' as it arbitrarily schematized, in the 3' to 5' direction on RNA but not DNA (Jia et al., 2012; Pyle, 2008). However, this translocation can displace an annealed DNA or RNA 'top strand' oligo with approximately equal efficiency (Lim et al., 2017 and Puno MR, unpublished). We sought to take advantage of this by employing a bottom strand (γ strand) in our substrate that is an RNA/DNA chimera (Figure 36). We reasoned that Mtr4 would be able to translocate along the RNA portion until it hit the RNA/DNA junction, at which point it would hopefully remain stalled. We included a short RNA oligo annealed to the DNA portion of the chimera (α strand) to ensure that it would form an A-form helix and to aid in stalling the helicase, and a strand of DNA annealed to the RNA portion (β strand) to limit the 3' overhang to the 8 nt poly(A) tract. Heat-cooling these three strands and gel purifying the product forms the tri-partite substrates (αβγ) in Figure 36.

To estimate the length of RNA required to reach Dis3 through Mtr4 and the central channel, we used RNA lengths from known structures of $Exo10^{Dis3}$ bound to a stem-loop RNA (Makino et al., 2013b) (30 nt), $Exo12^{Dis3/Rrp6/Rrp47}$ (Makino et al., 2015) (~10-12 nt), and Mtr4 bound to RNA (Weir et al., 2010) (+5 nt). This gave us an approximate length of the 3' overhang after unwinding of 46 nt, which is in good agreement with data from RNase protection assays using *S. cerevisiae* Exo14 (Falk et al., 2017).

To see if this substrate was behaving as intended, ultraviolet (UV)-crosslinking was performed by using substrates of varying lengths with a 4-thio uracil (4SU) nucleobase two nucleotides from the 3' end of the γ strand (Figure 37A). Irradiating a sample of protein mixed with a 4SU containing oligonucleotide with long-wavelength

UV light allows the nucleobase to crosslink to proteins, so long as they are bound near the 4SU modification. Protein-nucleic acid adducts can be visualized after separation by SDS-PAGE via the fluorescein label in the 4SU containing bottom strand (Figure 37A).

We irradiated Exo13 and Exo14 complexes (nuclease-inactive) after prolonged incubation with substrates and either ATP or AMP-PNP. In the presence of either nucleotide Exo13 shows inefficient crosslinking to all the substrates tested, with weak signals that correspond to Dis3 and Rrp6 proteins. In contrast, Exo14 shows a strong signal corresponding to a crosslink to Mtr4 in all cases, leading to the speculative model depicted in Figure 37B. For Exo14 with ATP, use of the two shorter substrates generates signal corresponding to crosslinks to Exo9 subunits as well as Rrp6. Curiously, for the substrate with a variable region of 20 nt (28 nt overhang after unwinding when the poly(A) tract is accounted for), a crosslink to Dis3 appears. Three possible explanations that could account for this observation are 1. Mtr4 is displacing Rrp6 from the Exo9, as previously observed with a stem loop plus poly(U) RNA bound to Dis3 (Makino et al., 2015) and the RNA is highly extended 2. after the long incubation, unwound substrate is released from Mtr4 and engages Dis3 through the central channel as in previous structures (Kowalinski et al., 2016; Liu et al., 2016; Makino et al., 2013b) or 3. Dis3 is engaging the unwound RNA independent of the central channel. When using Exo14, ATP and the longer substrates, crosslinks to Dis3 dominate the reaction. These data lead to the (once again speculative) model presented in Figure 37C, and show that the longer substrates can thread through the core and engage Dis3 when ATP and Mtr4 are present.

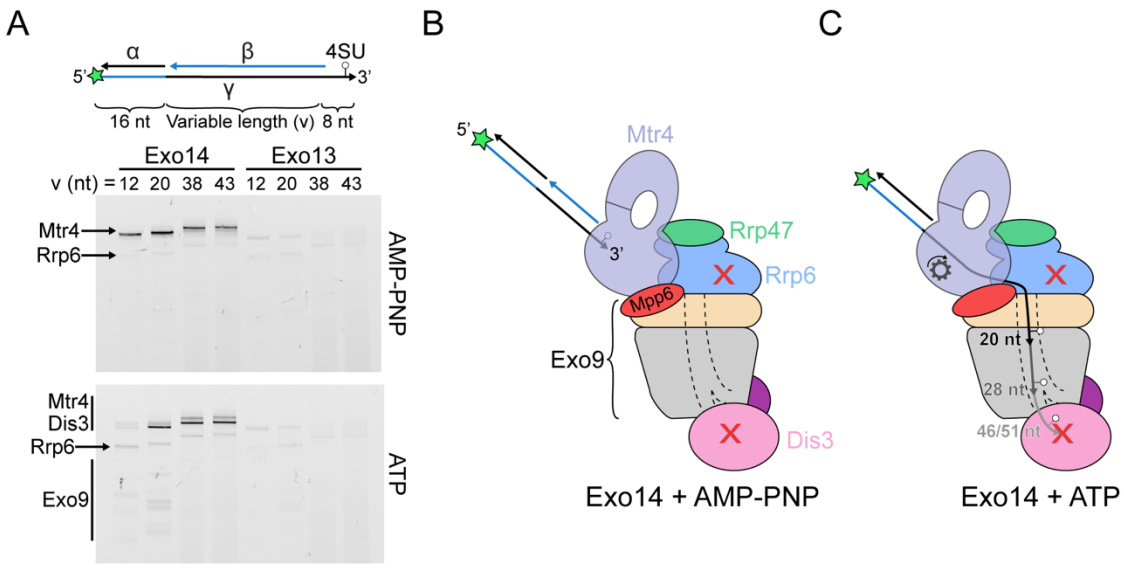


Figure 37. RNA crosslinking reveals ATP-induced substrate translocation by the Mtr4-exosome. (A) Top: Schematic representation of crosslinking substrates. Black represents RNA regions, blue represents DNA, green indicates the 5' fluorescein label on the γ strand, the marker indicates a 4SU modification 2 nt before the 3' end of the γ strand, and arrows point towards the 3' end. Bottom: SDS-PAGE analysis of RNA crosslinking to exosome complexes in the presence of ATP or AMP-PNP. Gels were imaged for fluorescein signal and crosslinked adducts are indicated. Data generated by Dr. M. Rhyan Puno with Exo13 and Exo14 supplied by John Zinder. (B and C) speculative models from the Exo14/AMP-PNP (B) or Exo14/ATP (C) cases in panel A. Inactive exonuclease subunits are indicated with red X's and helicase direction is shown with a gear.

The nuclear RNA exosome stimulates unwinding by Mtr4

Assured that the 3' end of our substrate was being fed to Dis3, we sought to directly observe unwinding. To this end, we performed helicase assays using αβγ with variable region v = 38 nt containing a 3' label on the β strand (Figure 38A). Given the high melting temperature of our substrate, reannealing of the displaced strand during workup precluded any activity measurement except when an unlabeled DNA oligo complementary to the β strand was included in the reaction (DNA Trap, see next section). Mtr4 alone was unable to unwind the substrate under the conditions employed in the assay, consistent with the low level of activity previously observed (Jia et al., 2012). Exo13$^{Dis3/Rrp6/Rrp47/Mpp6}$, having no intrinsic helicase activity, was also unable to unwind the substrate. Combining Mtr4 with Exo13 or Exo11$^{Dis3/Mpp6}$, both of which contain elements that recruit Mtr4 (Figure 29), produced a robust, ATP-dependent release of the β strand.

When using complexes with catalytically active Dis3 and/or Rrp6 on an RNA variant of this substrate, we observe similar trends to those seen in Figure 32 with respect to Mtr4- and ATP-dependent Dis3 activity (Figure 38B). These experiments showed that the nuclear exosome complex can stimulate Mtr4's RNA helicase activity for degradation of structured substrates and gave us confidence that we would be able to load Exo14 and engage Dis3 with this tripartite substrate.

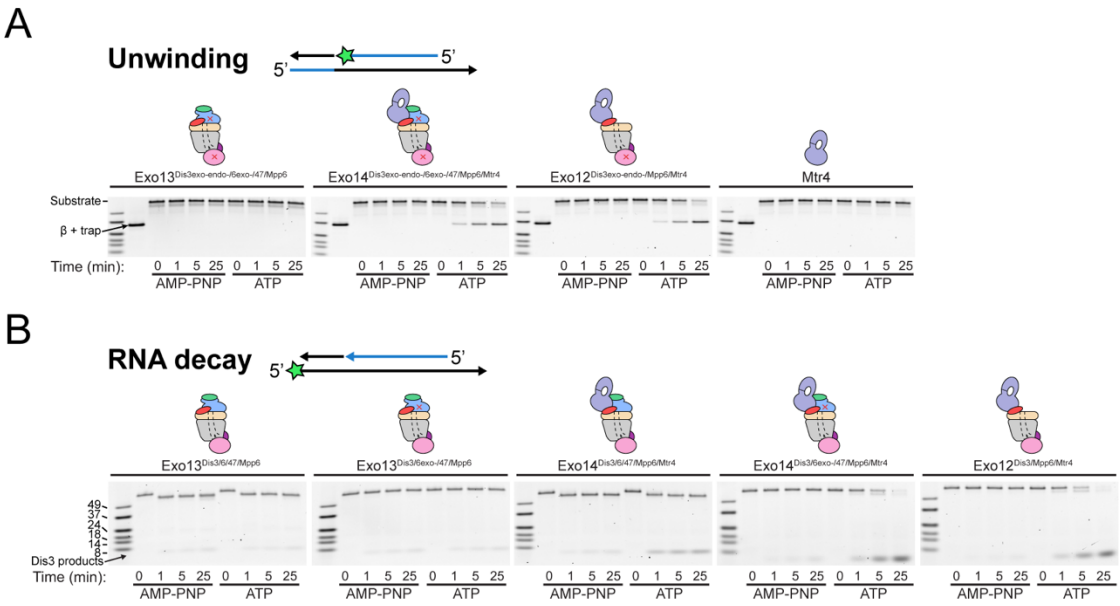


Figure 38. The exosome stimulates Mtr4's helicase activity for degradation of structured substrates. Gels are 4-20% acrylamide-TBE and were imaged for fluorescein signal. The substrate contains a variable region (v) of 38 nt as described in Figure 36. (A) RNA helicase assay. Top: schematic of the substrate used in the helicase assay. Black indicates RNA regions, blue is DNA, and the 5' fluorescein label is represented by a green star. Bottom: RNA unwinding time courses by Mtr4 in the presence or absence of catalytically inactive exosome complexes. Reactions were conducted using ATP or AMP-PNP as indicated. Markers in the first lane are (from top) 49, 37, 24, 18, 14, and 8 nt 5' fluorescein poly(A) RNA. A marker indicating the trap annealed to the released strand is included in the second lane of the gels. (B) RNA decay assay. Top: schematic of the substrate used in the decay assay. Schematics are as in panel A. Bottom: RNA decay time courses using the indicated complexes. Catalytically inactive Rrp6 is indicated with a red X. Reactions were conducted using ATP or AMP-PNP as indicated.

Loading Exo14 with a substrate stabilizes Mtr4 on the exosome and requires a trap for the displaced strand

As our final goal was to characterize the loaded Exo14 complex by cryo-EM, we first set out to optimize the solubility of the complex. Specifically, we were concerned that the C-termini of Rrp6 and Rrp47, both of which are predicted to be disordered and decrease the expression levels and solubility of the heterodimer (Costello et al., 2011; Wasmuth and Lima, 2017), would cause aggregation during concentration or upon vitrification. To solve this, we reconstituted 13-component complexes lacking the C-terminal 100 amino acids of Rrp6 and 80 amino acids of Rrp47, which will henceforth be referred to as Exo13tr. Amino acids 700-721 from Rrp6's sequence were appended onto the truncated Rrp6, as that stretch had previously been shown to increase activity of the enzyme without compromising solubility (see Materials and Methods) (Wasmuth and Lima, 2017).

The Exo13tr complex's activity was compared to WT (both with catalytically inert Rrp6) in a helicase-dependent RNA decay assay as previously (see Chapter 2), and showed nearly identical activity (Figure 39). Furthermore, when channel occlusion mutant of Rrp41 (see Chapter 1) was included in the complex, Exo13tr lost helicase-dependent Dis3 activity altogether, indicating that the substrate was entering the central channel of un-occluded Exo13tr with the help of Mtr4 and ATP. This result is inconsistent with a model in the literature wherein Mtr4 can feed RNA to Dis3's direct-access conformation (Delan-Forino et al., 2017), however that model relied on the unconfirmed hypothesis that a mutation in the S1 domain of Dis3 selectively disrupted direct-access RNA degradation by Dis3 (see Chapter 1, Discussion). Both Exo13 and Exo13tr were able to degrade poly(A) RNA in the presence of AMP-PNP but, interestingly, this activity was enhanced when using ATP.

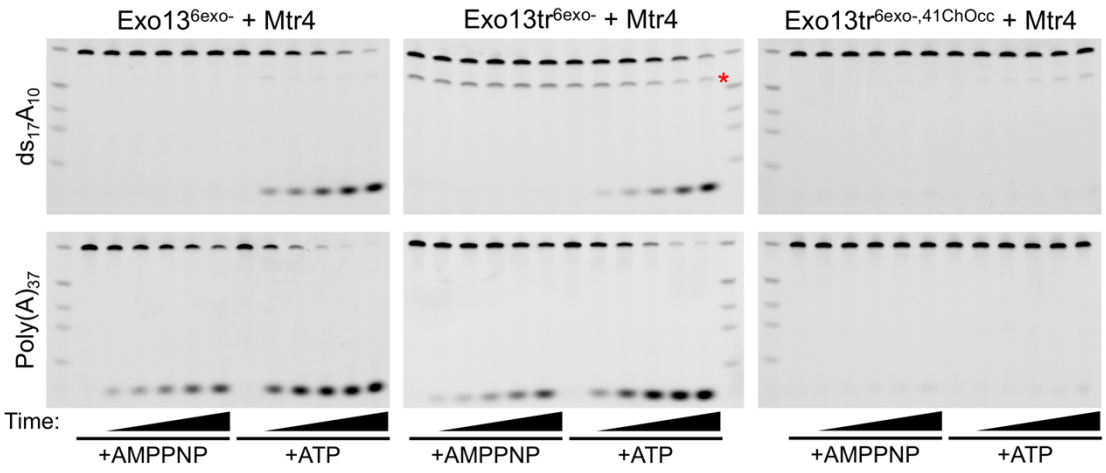


Figure 39. Exo13 containing truncations in Rrp6 and Rrp47 retains near-WT activity. RNA decay assays using the indicated complexes on 37 nt poly(A) and $ds_{17}A_{10}$ RNA (see Chapter 2, Materials and Methods), both of which contain 5' fluorescein labels. Exo13tr refers to a complex containing Rrp6 and Rrp47 C-terminal truncations described in the text. Rrp6 contains a D238N mutation to render it catalytically inactive in all cases. The red asterisk marks an artifact from workup of that particular reaction. Gels are 15% acrylamide TBE-urea and imaged for fluorescein signal. Markers are (from top) 37, 24, 18, 14, and 8 nt 5' fluorescein poly(A) RNA. The Dis3 product runs at the bottom of the gel.

A necessary but absent control would be to test decay on 37 nt poly(A) RNA in the absence of Mtr4, but nevertheless these data suggest that Mtr4 is competing with Dis3 for the 3' end of this particular substrate when it cannot translocate.

Next, we sought to optimize conditions for preparative loading of the substrate into an exonuclease dead complex. Given that the previous Exo14 reconstitution protocol (Figure 29) required a buffer that was optimized for Mtr4:exosome stoichiometry and not yield (see Chapter 2 Materials and Methods), we opted to mix 1.1-fold excess Mtr4 with Exo13tr prior to addition of RNA rather than use reconstituted Exo14. After mixing the two protein components and incubating on ice, the proteins were added to RNA substrates and incubated at room temperature. Then, either AMP-PNP (Reaction 3), ATP (Reaction 4), ATP plus a DNA Trap complementary to the β strand (Reaction 5), or ATP plus RPA, a trimeric protein complex from *S. cerevisiae* that binds ssDNA (courtesy of Dr. Laurent Cappadocia, Reaction 6) was added (Figure 40). After incubation at room temperature, the reactions were fractionated by gel filtration (Figure 40B) and peak fractions were analyzed the for protein and nucleic acid content by PAGE with SYPRO Ruby and SYBR Gold staining, respectively (Figures 40C, 40D, and 40E). As controls, substrate only (Reaction 1) or Exo14 + ATP with no added nucleic acid (Reaction 2) were injected.

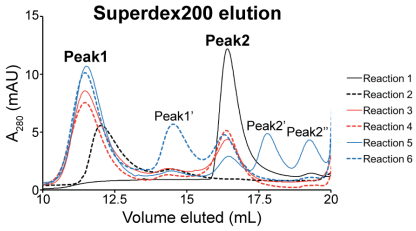
For the substrate alone (Reaction 1, solid black curve), a single peak centered at 16.4 mL containing nucleic acid and no protein (Peak 2) was observed as expected (Figure 40). The UV traces for the Exo14 complex alone (Reaction 2, dashed black curve) showed a broad peak centered at 12.1 mL (Peak 1), with a second minor peak centered at 14.5 mL, corresponding to the exosome and Mtr4, respectively (Figure 40B). Gels of the first peak showed that it contained Exo13tr components but sub-stoichiometric amounts of Mtr4 and no RNA (Figures 40C and 40D).

Figure 40. Gel filtration analysis of Mtr4-exosome loading reactions. (A) Table describing the shorthand for the reaction conditions used. 'RNA' refers to an αβγ tripartite substrate with v = 20 nt. (B) A_{280} traces from gel filtration (Superdex200, 24 mL column volume) from the six reactions. (C) SYPRO Ruby stained SDS-PAGE of the load and first peak from each reaction. Three 0.5 mL fractions that approximately encompassed the peak's centroid were pooled for each 'Peak' sample throughout this figure. For the Reaction 6 load, RPA subunits (Rfa1, Rfa2, and Rfa3) are highlighted. Asterisks in the Reaction 3 Peak lane indicate unknown protein contamination. (D) Nucleic acid gels of the load and two main peaks. Samples were treated with Proteinase K prior to separation. Gels are 4-20% acrylamide-TBE and were visualized with SYBR Gold for detection of nucleic acid. (E) High-resolution gel of Peak 1 from Reaction 5. The sample from panel D was run in both lanes 1 and 4 alongside standard curves of the indicated markers. Gel is 4-20% acrylamide-TBE, but was run for 50% longer than the gels in panel D. Visualization was as in panel D. (F) Nucleic acid gels of miscellaneous peaks from Reactions 5 and 6. Sample preparation and gel are as in panel D.

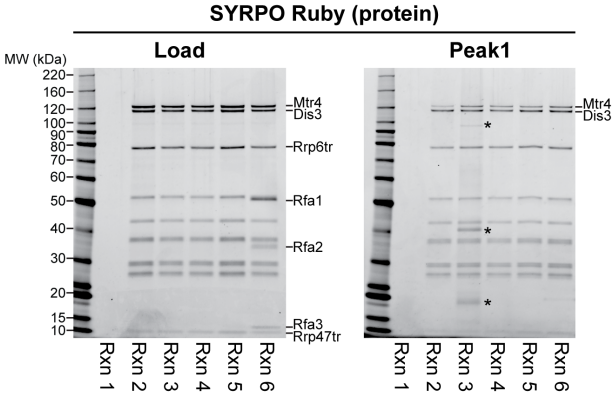
A

Reaction:	1	2	3	4	5	6
RNA	+	-	+	+	+	+
Exo14	-	+	+	+	+	+
ATP	-	+	-	+	+	+
AMP-PNP	-	-	+	-	-	-
DNA Trap	-	-	-	-	+	-
RPA	-	-	-	-	-	+

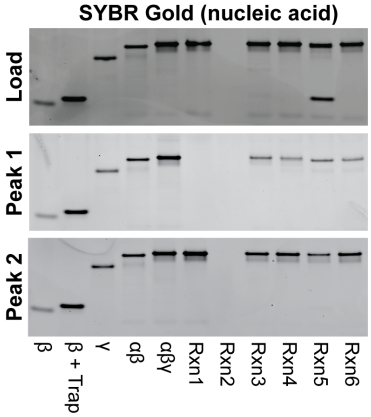
B



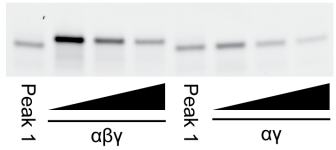
C



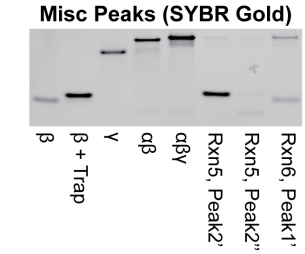
D



E



F



While this result contrasts the result from gel filtration in Chapter 2 (Figure 29), which showed stoichiometric association of Mtr4 with Exo13, the reactions presented here were conducted at much lower concentrations and using different reaction and gel-filtration buffers. For the remaining reactions, Peak 1 had a higher UV absorbance, was narrower, and eluted earlier (centered at ~11.5 mL) than with protein alone. Reactions 3 and 4 (red curves in Figure 40B) behaved nearly identically, with greater amounts of Mtr4 present in the protein gel of Peak 1 (Figure 40C) than for without substrate (Reaction 2), and mostly αβγ, or the full substrate, in Peak 1 (Figures 40D). For Reaction 4, the lack of a significant proportion of αγ (the expected product of unwinding) in Peak 1 suggests that the β-strand is either not being displaced or is co-migrating with the complex throughout gel filtration and re-annealing during the workup to make the sample for the gel. Reactions 5 and 6 (blue curves in Figure 40B), by contrast, both contain traps to sequester the displaced strand and show αγ and no αβγ in their Peak 1 (Figures 40D and 40E). For Reaction 5, two additional peaks appear: Peak 2' contains the trapped β-strand and Peak 2'' contains the excess DNA trap oligo, though it is difficult to visualize, potentially because it is C-rich and SYBR gold has poor fluorescence when bound to poly(dC) (Tuma et al., 1999) (Figure 40F). Reaction 6 has an additional peak (Peak 1') that contains the β strand (presumably bound to the RPA trimer), though it is also is contaminated with αβγ from Peak 2 (Figure 40F). These experiments demonstrate that nucleic acid strengthens the Mtr4-exosome interaction independent of substrate loading (compare Reaction 3 to Reactions 4 through 6), and that the observation of efficient loading is aided by use of a DNA or protein trap.

Crosslinking of Exo14 to stabilize the complex

3D reconstructions from preliminary cryo-EM analysis of substrate loaded Exo14tr showed that some of the subunits were substoichiometric or entirely missing. Among these, Csl4 and the Rrp6 EAR were the most noticeably weak. To address this, we optimized protocols to crosslink the complex using homo-bifunctional amine-to-amine (glutaraldehyde, ~5 Ångstrom spacer) and sulfhydryl-to-sulfhydryl ($BM(PEG)_3$, ~18 Ångstrom spacer) crosslinking molecules (Figure 41).

Treating the substrate-loaded Exo14tr complex with increasing amounts of glutaraldehyde resulted in mobility shifts in many of the subunits, most noticeably Rrp6 and Mtr4, and generated a smear of high molecular weight species (Figure 41A, left half). By contrast, treatment with $BM(PEG)_3$ resulted in discrete high molecular weight bands, with few noticeable shifts among the majority of subunits (Figure 41A, right half), presumably because surface cysteines are scarce relative to surface lysines. The most obvious effect was the disappearance of the Rrp6 and Csl4 proteins, and the strengthening of signal around the Mtr4 band. We suspected that the signal that co-migrated with Mtr4 was derived from an Rrp6/Csl4 crosslinked species, and performed Western blotting analysis on the crosslinked and un-crosslinked complex to investigate (Figure 41B). When the Rrp6 and Mtr4 blots were aligned to the direct blue protein stain, a shift in molecular weight of the Rrp6 signal to approximately the same mobility as the un-crosslinked Mtr4 band was observed, supporting our suspicion. Additionally, ~50% of the Mtr4 signal shifted upwards, indicating that its cysteines form crosslinks to other cysteines in the complex, though this cannot be rationalized using current structural models.

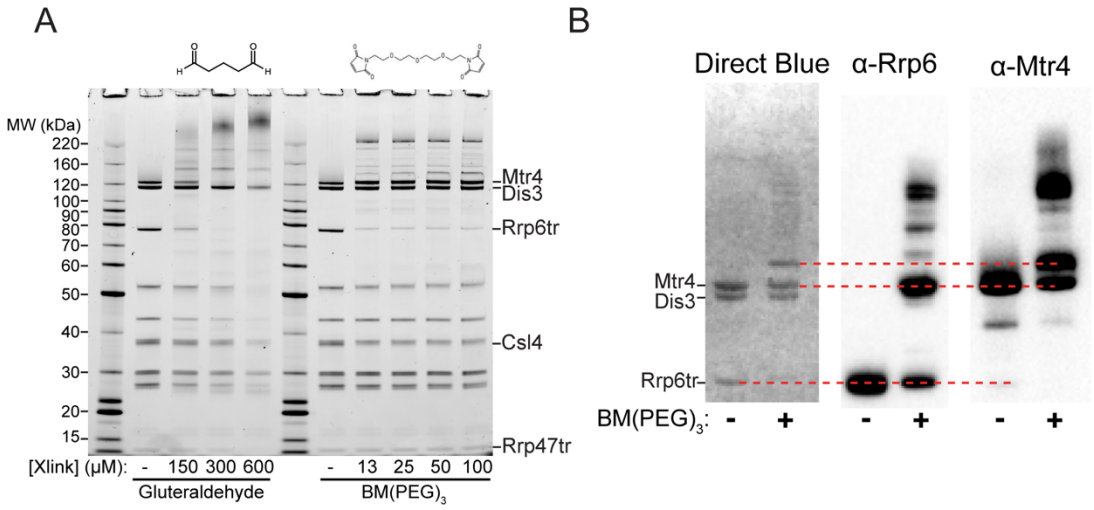


Figure 41. Protein-protein crosslinking using glutaraldehyde or BM(PEG)$_3$. (A) SYPRO Ruby stained SDS-PAGE showing Exo14tr complexes treated with different concentrations of an amine-to-amine (glutaraldehyde) or sulfhydryl-to-sulfhydryl (BM(PEG)$_3$) crosslinking molecule. Molecular structures are shown above the gel. (B) Western blot analysis of crosslinked (50 μM BM(PEG)$_3$) and un-crosslinked Exo14tr. The same PVDF membrane was visualized using protein stain (left, direct blue) or a Western blot and manually aligned. Rabbit polyclonal antibodies were raised against Rrp6 residues 129-518 or Mtr4 residues 80-1073 and affinity purified (see Chapter 2, Materials and Methods). Western blots were treated with a secondary antibody conjugated to HRP and visualized for chemiluminescence. Membrane was stripped and re-probed after Rrp6 detection using standard protocols.

S. cerevisiae Rrp6 contains three cysteines in its entire sequence, one of which, C553, is 12 or 16 Ångstroms away from C18 and C66 of Csl4, respectively (Figure 42A). The other two are either buried (C252) or out of the way of the majority of proteins (C101) in the N-terminal 'lid' (Figure 42A). An Exo13tr complex containing $Rrp6^{C553S}$ was reconstituted and treated with $BM(PEG)_3$ in parallel with WT. Crosslinks were observed by Western blotting and SYPRO Ruby staining (Figures 42B and 42C). $Rrp6^{C553S}$ showed weaker crosslinking by Western blot, and fewer high-molecular weight species were observed in the corresponding SYRPO stained gel, indicating this residue is participating in cysteine-to-cysteine crosslinking in the complex. In both the Western and the SYPRO it is clear that un-crosslinked Rrp6 intensity of the mutant diminishes in the presence of the $BM(PEG)_3$ (though not as dramatically as in the WT case), which suggests that the two other cysteines may participate in yet unknown protein-protein interactions.

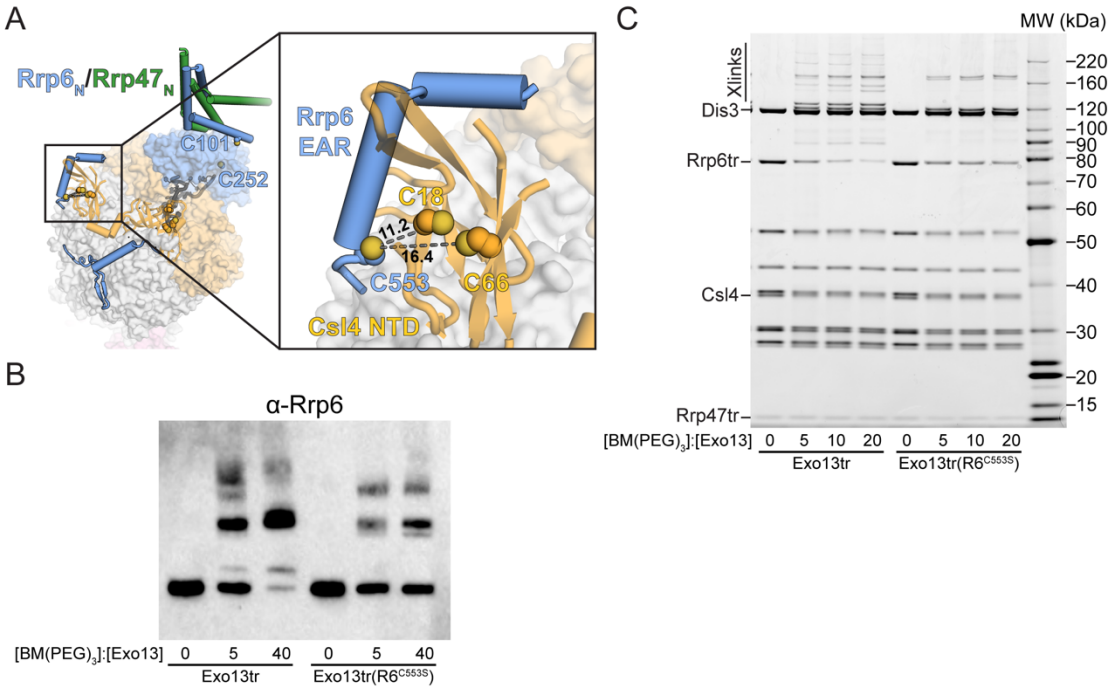


Figure 42. Structural modeling of Rrp6 crosslinking. (A) Location on the nuclear exosome of Rrp6's three cysteines. Model is from PDB 5K36, with Rrp6/Rrp47 N-terminal bundle from PDB 5C0W. Distances in zoom panel are shown in Ångstroms. Panel generated using Pymol (Schrödinger). (B) Western blot of Exo13tr containing WT and C553S mutant Rrp6 treated with different amounts of $BM(PEG)_3$. Antibody and visualization is as in previous figure. Reaction contained 500 nM Exo13tr. (C) SYPRO Ruby stained SDS-PAGE showing WT and mutant Exo13 complexes treated with different concentrations $BM(PEG)_3$. Reaction contained 500 nM Exo13tr.

Preparation of loaded ScExo14tr for cryo-EM data collection

Having optimized the loading reaction and protein crosslinking, we set out to collect cryo-EM data for high resolution structure determination of the loaded Exo14tr complex. The substrate in Figure 36A with v = 38 nt was chosen for these studies, as it showed strong 3' UV-crosslinks to Dis3 after ATP addition (Figure 37A) and could reach Dis3 for degradation with ATP and Mtr4 present (Figure 38B). Rather than purifying the loaded complex via gel filtration, which would risk dissociating the loaded complex both through dilution on the column and because it is time consuming, the loading reaction was simply concentrated and desalted into reaction buffer containing ADP after incubation. Furthermore, the components that gel filtration removes, namely excess Mtr4, substrate, and unreacted ATP, are small enough that they will either not be seen or can be easily removed by 2D or 3D classification during data processing. ADP was added to the final buffer because Mtr4 has previously been crystallized bound to RNA and ADP (Weir et al., 2010) and because addition of ADP enhanced Mtr4's ability to shield RNA when bound to the exosome in RNase protection assays (Falk et al., 2017). After desalting, the complex was diluted to a pre-determined concentration, treated with $BM(PEG)_3$, quenched with DTT, adjusted to the final detergent concentration, and grids were prepared.

Data collection, processing, and 3D classification

We collected movies using a target defocus range of -1.5 to -2.5 μm on a Titan Krios equipped with K2 Summit direct detector over a span of approximately 40 hours. The complex appears to have a desirable behavior (in terms of particle size and distribution) in relatively thick ice (Figure 43A), but not when then ice is excessively thick. Because of this, holes for collection were picked based on an ideal value for pixel density (which relates to ice thickness) that was determined during screening

using a python script (Laurent Cappadocia, unpublished). This resulted in 1483 movies, which after motion correction (Zheng et al., 2017) and CTF estimation (Zhang, 2016) was reduced to 1144 using a 4.4 Ångstrom max resolution cutoff plus manual pruning. Auto-picking in Relion 2 (Kimanius et al., 2016) resulted in an initial batch of 245,000 particles, 222,625 of which remained after conservative selecting of 2D-class averages in CryoSPARC (Punjani et al., 2017) to remove debris and other aberrant particles (Figure 44). From this batch, an initial 3D classification in Relion resulted in 6 classes (Figure 43B), all of which resemble various subcomplexes of the exosome in size and shape but differ in their resolution, population, and apparent stoichiometry of the complex. The two most populated classes (Classes 4 and 6) from this were pooled and subjected to further 3D classification (Figure 44).

A common feature of the six 3D classes at this point was that they all contained the non-catalytic core (the most clear element being the C-terminal helix of Rrp45, highlighted in Figure 43B with a red circle) and Dis3 in the through-channel conformation (Liu et al., 2014; 2016; Makino et al., 2013b). The pooled classes resulted in a volume that encompassed these elements as well as additional densities above the S1/KH subunit Rrp4, which we will refer to as 'top densities'. These densities were poorly defined relative to the core region, so focused classification was performed to separate particles that contained these densities (42%) from particles that did not (58%) (see Materials and Methods) (Figure 44). The resulting classes were subjected to a second round of focused classification for Csl4, which was relatively weak in both classes and throughout. After this point, the class that contained top densities was separated into two 3D classes, Class I and Class II (Figures 44, 45A and 45B). Attempts at further classification did not result in significantly different structures. Class II additionally contains densities on the side of the exosome proximal to the Rrp4 S1/KH protein and Rrp42 PH-like protein as well as extra densities above

the S1/KH ring (Figures 44 and 45B). Finally, focused classification for Csl4 from the particles that did not contain top densities resulted in Class III (Figures 44 and 45C). Analysis of particle orientation from 3D refinement in CryoSPARC revealed orientation bias in all three classes, though this was more severe for Classes I and II than for Class III (Figure 44).

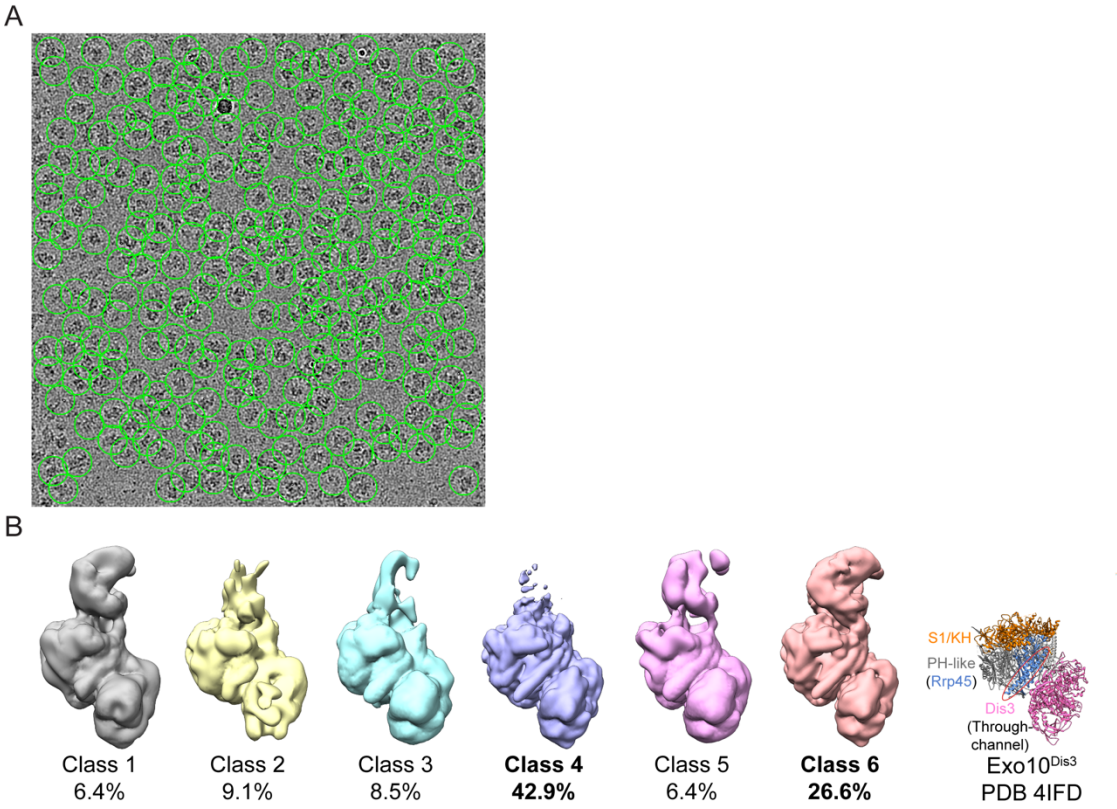
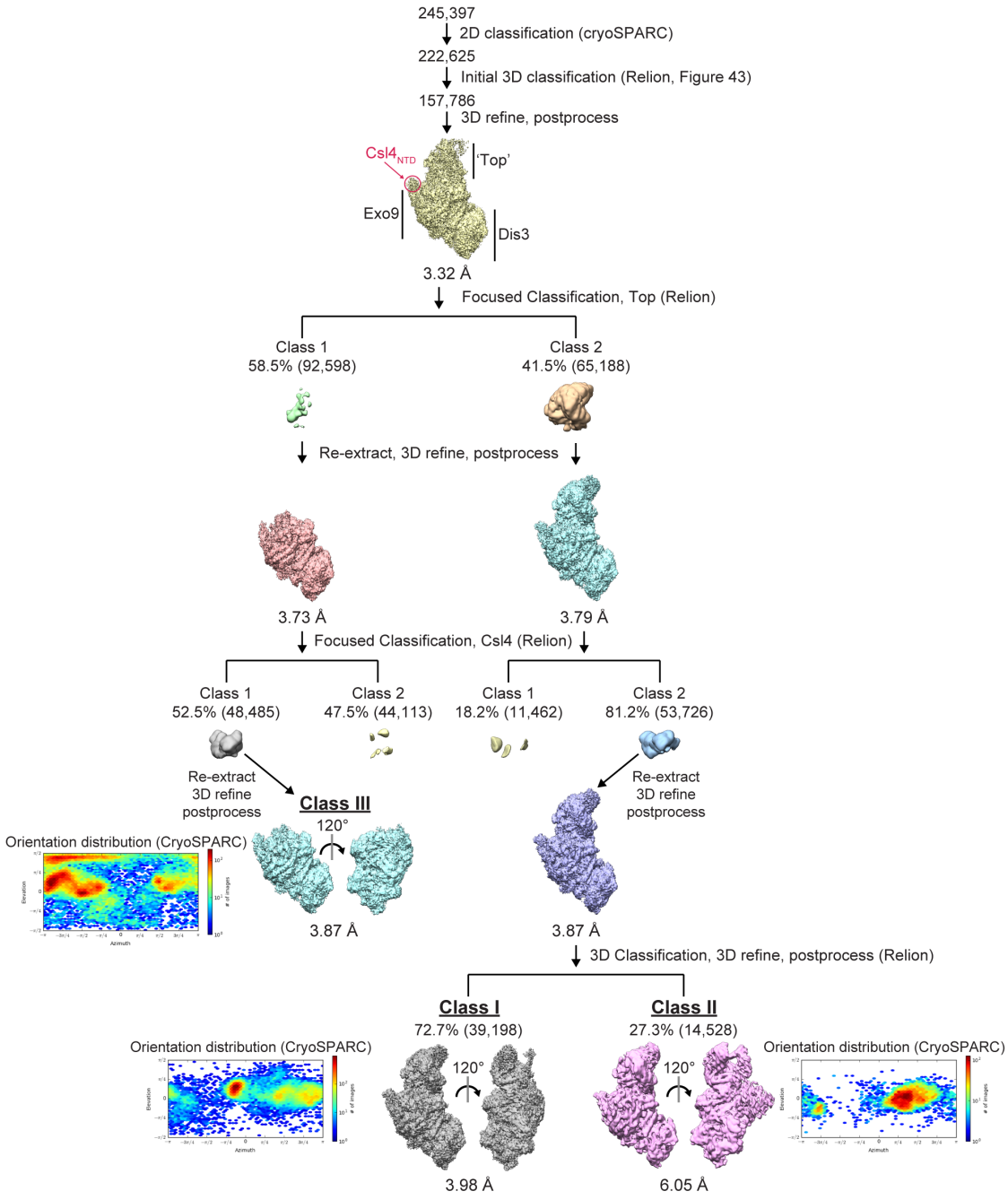


Figure 43. Initial characterization of a cryo-EM dataset. (A) Representative micrograph used in analysis. Image was low-pass filtered at 20 Å. Green circles are 250 Å in diameter and show the un-pruned output from auto-picking in Relion. (B) Initial 3D classification from 222,625 particles in Relion. The classes that were used for subsequent analysis (Classes 4 and 6) are highlighted in bold. $Exo10^{Dis3}$ from PDB 4IFD (with RNA and Rrp6 EAR removed) is shown for comparison, with the C-terminal helix of Rrp45 highlighted by a red oval. Figure generated using Chimera (Pettersen et al,. 2004).

Figure 44. Summary of 3D classification. Particle count and resolution are shown at different steps of 3D classification in Relion. The three classes used in subsequent analyses (Class I, Class II, Class III) are highlighted in bold font. For post-processed 3D volumes, masked densities are shown. Next to each of these classes the orientation distribution generated by a homogenous 3D refinement of those particles in CryoSPARC is shown. Each point represents a discrete orientation using a horizontal coordinate system, and the color of a point represents how many particles are in that particular orientation for the 3D reconstruction. White regions indicate that there are no particles in those orientations.


245,397
2D classification (cryoSPARC)
222,625
Initial 3D classification (Relion, Figure 43)
157,786
3D refine, postprocess
Csl4NTD
'Top'
Exo9
Dis3
3.32 Å
Focused Classification, Top (Relion)
Class 1
58.5% (92,598)
Class 2
41.5% (65,188)
Re-extract, 3D refine, postprocess
3.73 Å
3.79 Å
Focused Classification, Csl4 (Relion)
Class 1
52.5% (48,485)
Class 2
47.5% (44,113)
Class 1
18.2% (11,462)
Class 2
81.2% (53,726)
Re-extract
3D refine
postprocess
Re-extract
3D refine
postprocess
Class III
120°
Orientation distribution (CryoSPARC)
3.87 Å
3.87 Å
3D Classification, 3D refine, postprocess (Relion)
Class I
72.7% (39,198)
Class II
27.3% (14,528)
Orientation distribution (CryoSPARC)
120°
120°
Orientation distribution (CryoSPARC)
3.98 Å
6.05 Å

Structural modeling into 3D-classes

Because of the relatively low resolution of the data, rudimentary docking of pre-existing crystal structures in Chimera was performed (Pettersen et al., 2004) without significant building or refinement (Figure 45). Elements that were clearly not encompassed by the volumes after docking were removed, resulting in the models presented in Figures 45 through 48. The crystal structure of the yeast exosome with Dis3 bound to RNA in the through-channel conformation (Makino et al., 2013b) (PDB 4IFD) could be readily docked into all three classes, though density for the Rrp6 EAR from that structure was weak or absent (Figures 45, 46C, and 46E). Mtr4 fit well into the top density for Classes I and II, including the bound RNA and ADP from PDB 2XGJ (Weir et al., 2010) (Figures 45, 46A, 46B, and 46E), though density for the arch was comparatively weak in Class I (Figures 45A and 45B). In these classes, RNA appears to continue from the 3' end of the RNA in Mtr4 through the central channel, and there is weak but contiguous density connecting it to the Dis3 RNA (Figures 46B, 46D, and 46E). Densities extending from the 5' end of the RNA was weak or absent in both Mtr4-containing classes, and the duplex region of the substrate was not apparent. Class III also appears to contain RNA throughout its central channel engaging Dis3 (Figure 46F). The catalytic module of Rrp6 and the N-terminal helical bundle with Rrp47 could not be modeled into any of the three classes. Mpp6 bound to Rrp40 was also not observable in any of the three models, though some additional densities in that region were present at high contour levels. While it is likely absent, it remains possible that Mpp6 is either present in multiple conformations or in a subset of the particles included in the classes.

Interestingly, Mtr4 sits on the same conserved surface of Rrp4 in these models as Rrp6 does in previous structures (Wasmuth et al., 2014; Zinder et al., 2016) (Figure 47). Another notable aspect of the models is that the Trf4 and Air2 peptide motifs that

bind the Mtr4 catalytic core could accommodated (Falk et al., 2014; Losh et al., 2015) (Figure 48), raising the possibility that TRAMP may engage the exosome in a similar manner as we observe for Mtr4 alone.

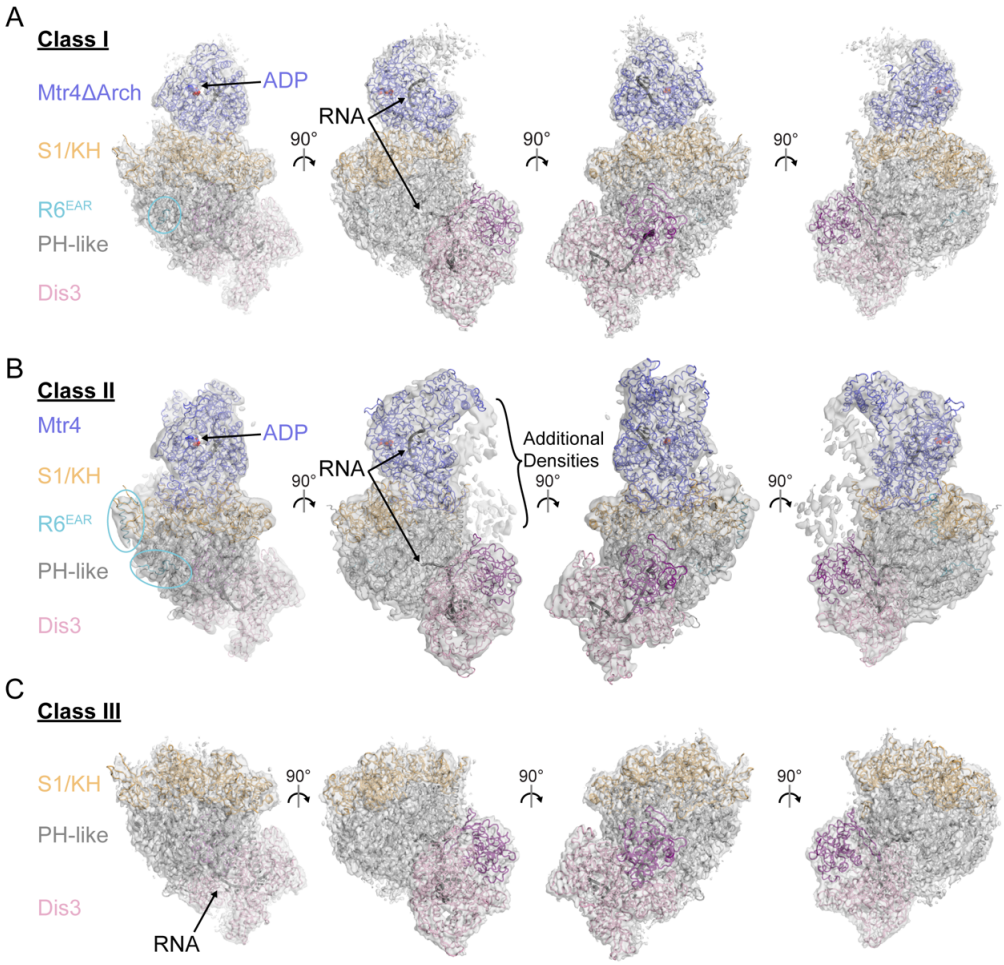
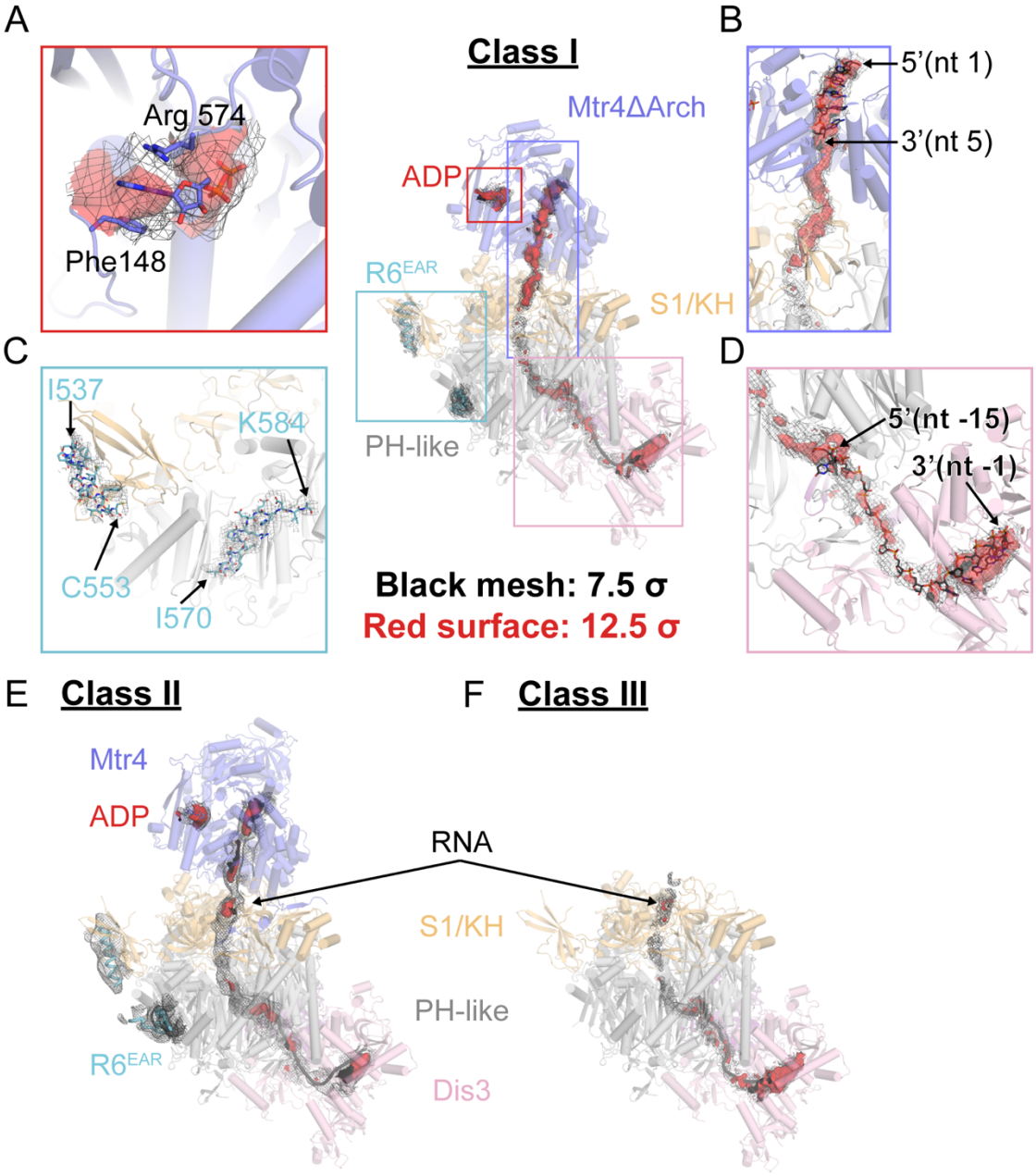


Figure 45. Docking of crystal structures into 3D classes. Grey surface is the post-processed, masked density contoured to 7.5 (arb. units) throughout. Through-channel Dis3 bound to RNA, Exo9 and Rrp6 EAR (if present) are from PDB 4IFD. Mtr4 bound to RNA and ADP is from PDB 2XGJ. Figure was generated using Pymol (Schrödinger). (A) Docking into Class I. RNA is indicated as well as a fragment of the Rrp6 EAR. Mtr4's arch domain, which spans residues 607 to 878, was removed from the model as well as most of the Rrp6 EAR. (B) Docking into Class II. RNA is indicated as well as a fragment of the Rrp6 EAR. Additional densities that could not be modeled into are highlighted. (C) Docking into Class III. RNA is indicated. The Rrp6 EAR from PDB 4IFD was not encompassed in the density and removed from the model.

Figure 46. RNA and other features in 3D classes. Grey mesh is the masked, post-processed density contoured to 7.5 σ and red surface is contoured to 12.5 σ throughout. Models are as in the previous figure, though an additional fragment of the Rrp6 EAR is shown for Class I for illustrative purposes. Figure generated using Pymol (Schrödinger). (A through F) Features in Class I. (A) ADP is the ATPase site of Mtr4. Sidechains of Phe148 and Arg574 that sandwich the adenosine moiety of ADP are highlighted. Both the mesh and surface were carved to within 3 Å of the ADP molecule. (B) RNA in and around Mtr4. 5' and 3' ends of the RNA from PDB 2XGJ are indicated. Both the mesh and surface were carved to within 3 Å of a place-holder RNA model. (C) Density for the Rrp6 EAR. Note that the lack of red surfaces in this panel is because no signal was observed at 12.5 σ. The amino acid boundaries of the model are indicated. Mesh was carved to within 3 Å of the Rrp6 EAR model. (D) RNA in and around Dis3. 5' and 3' ends of the RNA from PDB 4IFD are indicated. Both the mesh and surface were carved to within 3 Å of a dummy RNA model. (E) Density through the central channel, Dis3, and Mtr4 of Class II. Both the mesh and surface were carved to within 5 Å of a place-holder RNA model. Also shown are densities caved to within 5 Å of the ADP molecule in Mtr4 and the Rrp6 EAR. (F) Density through the central channel and Dis3 of Class III. Both the mesh and surface were carved to within 3 Å of a place-holder RNA model.


A
Arg 574
Phe148
Class I
Mtr4ΔArch
ADP
R6EAR
S1/KH
PH-like
B
5'(nt 1)
3'(nt 5)
C
I537
K584
C553
I570
D
5'(nt -15)
3'(nt -1)
Black mesh: 7.5 σ
Red surface: 12.5 σ
E Class II
Mtr4
ADP
RNA
S1/KH
PH-like
R6EAR
Dis3
F Class III

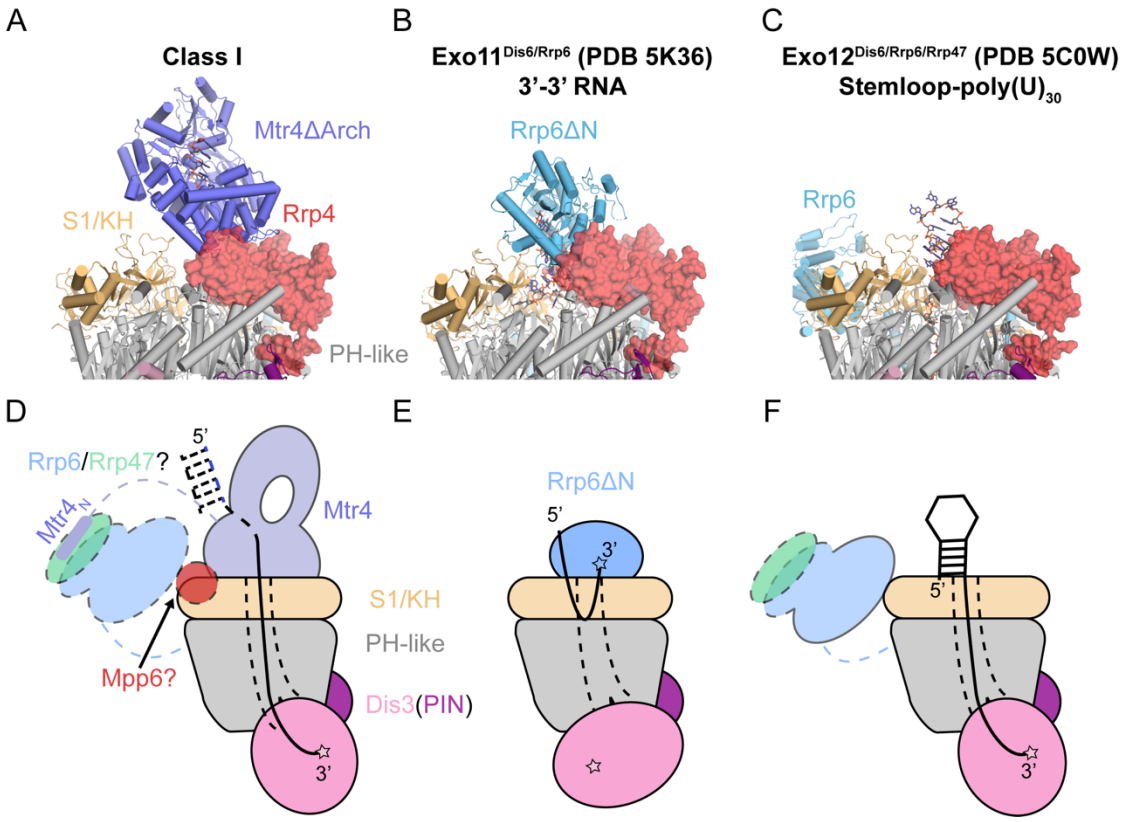


Figure 47. Rrp6 can be displaced from the S1/KH ring. Structures were aligned based in the Rrp4 chain. Panels A through C were generated using Pymol (Schrödinger). (A) Mtr4 sits on top of Rrp4 (shown as a red surface) in the model from Class I (and Class II, though that is not shown). (B) Rrp6's catalytic module sits on the same surface in a structure of Exo11$^{Dis3/Rrp6}$ bound to a 3'-3' RNA and other structures (not shown). (C) Rrp6's catalytic module is displaced from Rrp4 in a crystal structure where Dis3 is engaging a structured RNA with long overhang through the central channel. (D through F) Schematic representations of the above panels. RNA 5' and 3' ends are indicated, and dashed lines surrounding subunits or in nucleic acid indicates that they are disordered or missing in the corresponding structure. Dashed lines in the core represent the central channel.

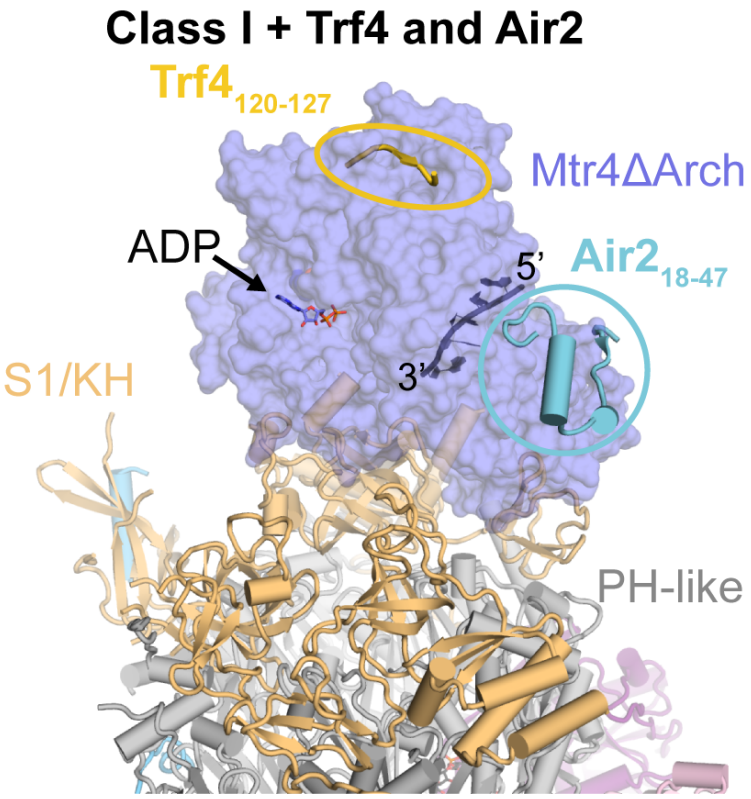


Figure 48. TRAMP subunits can be accommodated in the model from Class I. Structural model from the density in Class I. PDB 4U4C was aligned to the model, and the Trf4 and Air2 peptides contained within it are highlighted. Amino acid boundaries of Trf4 and Air2 peptides are indicated. Figure generated using Pymol (Schrödinger).

Discussion

These studies represent a proof of principle for employing a specialized mixed RNA/DNA substrate to structurally characterize helicase-dependent RNA decay by the *S. cerevisiae* exosome. Using this substrate, we observe a state of the complex in which Rrp6/Rrp47 has been displaced from the S1/KH cap by Mtr4 for channeling to Dis3. We envision this state as it relates to the competition for the 3' end of the RNA between Rrp6 trimming and Mtr4 channeling in helicase-dependent RNA decay (Figures 32, 38B and 39): when Mtr4 'wins' the competition, it can temporarily displace Rrp6/Rrp47 from the core and ensure it feeds the RNA along a path through the central channel to Dis3 (Figure 47D). This displacement of Rrp6 from the cap reminiscent of that observed in a crystal structure of $Exo12^{Dis3/Rrp6/Rrp47}$ bound to a stem-loop $poly(U)_{30}$ RNA (Makino et al., 2015) (Figures 47C and 47F), and may be an important facet of RNA decay by the nuclear exosome. Affinities between the conserved top surface of Rrp4 and Mtr4 or Rrp6's catalytic module are likely negligible, as neither binds the exosome without special elements to tether it (Schuch et al., 2014; Wasmuth and Lima, 2017; Wasmuth et al., 2014; 2017), so stable binding to this site by these enzymes is likely facilitated by RNA. Consistent with this, all high-resolution crystal structures of Rrp6 bound to the exosome to date include nucleic acid in its active site (Wasmuth et al., 2014; 2017; Zinder et al., 2016). After the RNA is degraded by Dis3, Mtr4 is no longer tethered to Rrp4 and can dissociate, potentially allowing Rrp6/Rrp47 to replace it. The interplay between competition for the RNA 3' end and competition for binding to Rrp4 between Mtr4 and Rrp6/Rrp47 may prove to be a fruitful area for future studies.

As previously mentioned, the catalytic module of Rrp6 and its N-terminal helical bundle could not be modeled into any of the final classes, while Rrp6's EAR

was weak or absent. A previous model from cryo-EM of the cytoplasmic exosome bound Ski7 observed a similar effect: Ski7's globular domain was entirely absent from 3D reconstructions while its EAR domain, which binds the same surface as Rrp6's, was weak relative to other regions (Liu et al., 2016). However, the investigators observed a signal with roughly the size and shape of Ski7's globular domains in aligned 2D class averages from negative stain EM (but not cryo-EM), though it was clear that it was not adopting a single conformation relative to the rest of the complex. This could also be the case for the Rrp6/47, and future studies may need to use negative stain EM or label Rrp6/Rrp47 with gold (or another electron-rich tag) in order to observe it. That being said, the density on the side of the exosome in Class II, while weak, vaguely resembles Rrp6's catalytic module. This class represents a small proportion (<10%) of the total particles, however, and may have only be stabilized by crosslinking.

Using the current strategy, classes containing Mtr4 represent approximately 40% of the total particles, though 20% of those lack Csl4. The Mtr4-containing class is further split 70% and 30% (into Classes I and II), leaving a small proportion of the desired complexes from the original population in spite of efforts to ensure stable association of Mtr4 through crosslinking and other protocols. Absence of Mpp6 as well as poor occupancy for Csl4 and thus Rrp6/Rrp47 (Rrp6 cannot associate with the exosome without Csl4) is at odds with their apparent stoichiometry as measured by SDS-PAGE and densitometry. This suggests that the complex is denaturing during grid preparation, or that populations of sub-stoichiometric complexes (which may represent a small fraction of the total population) are preferentially entering the holes in the grid over intact complexes. Conditions such as the crosslinking, reaction buffer, grid choice (graphene coating has been previously employed to resolve some of these issues), and detergent may need to be further optimized to mitigate these issues.

Furthermore, use of particle polishing and focused refinement (which may require larger datasets) may greatly enhance the resolution of the reconstructions and allow for more elegant structural modeling than simple docking.

These studies also provide a framework for the addition of more cofactors to the 14-component complex. With the determinants for and effects of Mtr4 recruitment more or less resolved, future studies can include the Trf/Air heterodimer to degrade substrates for which Mtr4's unwinding activity would not be sufficient. Additionally, reconstitution of co-transcriptional recruitment of the TRAMP-exosome by the Nrd1-Nab3-Sen1 (NNS) termination complex (Vasiljeva et al., 2008) provides a tantalizing target for use of biochemical and cryo-EM analysis in future work.

Materials and Methods

Oligonucleotides

All oligonucleotides were purchased and HPLC purified by Dharmacon. The sequence of the RNA top strand (α) used throughout this chapter is 5' AGCACCGUAAAGACGC 3' . Chimeric bottom strands (γ) for the tripartite substrates with v=12, 20, 38, and 43 are (from 5' to 3' end) dGdCdGdTdCdTdTdTdAdCdGdGdTdGdCdTCACCACACCACACAAAAAAAA, dGdCdGdTdCdTdTdTdAdCdGdGdTdGdCdTCACACCACACCACACCACACAAAAAAAA, dGdCdGdTdCdTdTdTdAdCdGdGdTdGdCdTCACACCACACCACACCACACCCACACCACCAAAAAAAA, and dGdCdGdTdCdTdTdTdAdCdGdGdTdGdCdTCACCACACCACACCACACCACACCACACCACACCACACCACACAAAAAAAA, where a 'd' preceding a letter indicates a deoxyribonucleoside. DNA blocking oligos (β) for v=12, 20, 38, and 43 are (from 5' to 3') dGdTdGdTdGdGdTdGdTdGdGdTdG, dGdTdGdTdGdGdTdGdTdGdGdTdGdTdGdGdTdGdT, dGdTdGdTdGdGdTdGdTdGdGdTdGdTdGdGdTdGdTdGdGdTdGdTdGdGdGdTdGdTdGdGdGdTdGdT, and dGdTdGdTdGdGdTdGdTdGdGdTdGdTdGdGdTdGdTdGdGdTdGdTdGdGdGdTdGdTdGdGdGdTdGdTdGdGdGdTdGdT. The sequence of the 'DNA Trap' used in Figures 38B, 40, and for the preparative loading reaction was 5' dAdCdAdCdCdAdCdAdCdCdAdCdAdCdCdAdCdAdC 3'. Substrates for UV crosslinking experiments in Figure 37A were ordered with a 5' fluorescein and a 4 thio-uridine 2 nt from the 3' end and were otherwise the same sequences as above. The substrate for the helicase assays in Figure 38A contained a 3' fluorescein label β-

strand and had a variable region length v = 38 nt. The substrate for the decay assay in Figure 38B contained a γ-strand that was of length v = 38 nt, was composed entirely of RNA and contained a 5' fluorescein label. Tripartite substrates were annealed by heat-cooling using excess α and β and purified by gel filtration on a Superdex200 increase 10/3000 GL column.

Cloning and expression of Rrp6tr/Rrp47tr

Rrp6 was amplified from a plasmid containing the full-length gene using 3' GTATGAGGATCCACTTCTGAAAATCCGGATGTACTTTTATC 3' and 5' CACCGTCGACTTACCTCCTCTTTTTAGCTGCCCTTGGTCCATTACTATCGCT AGATGATGGGTCGAATCTCCTTTTCTTCTTAATCTCCTCTAGCTTGACAGGA 3' as the forward and reverse primers respectively. The amplicon was cloned into the BamHI and SalI sites of pRSFduet-Smt3 (see Chapter 1, Materials and Methods). Then, Rrp47 was amplified from a vector containing the full-length protein using 5' GTATGACATATGGAAGATATCGAAAAGATAAAACCATA 3' and 5' AGACTCCTCGAGCTACTTCTTCCCTCCTTTCTTTTTTCC 3' and inserted into the NdeI and XhoI sites of that vector. Note that Rrp6 had to be cured of internal NdeI and BamHI sites prior to employing this strategy. The heterodimer was expressed in BL21 (DE3) RIL cells and purified by nickel, heparin, and gel filtration chromatographies similar to other exosome components (See Chapters 1 and 2, Materials and Methods).

Protein-nucleic acid UV crosslinking

Protein samples (200 nM) were incubated with 100 nM 5' fluorescein- and internal 4SU-labeled RNA in a buffer containing 20 mM Tris-HCl pH 7.0, 100 mM NaCl, 2.5 mM $MgCl_2$, 5 mM BME, and 2 mM AMPPNP or ATP. To induce

crosslinking, samples were exposed to long-range UV (365 nm) for 20 minutes at 4°C using a 4W handheld UV lamp (UVP). RNA-protein adducts were separated using NuPage 4-12% Bis-Tris protein gels (Thermo Fisher Scientific) and visualized using Typhoon FLA 9500 laser scanner (GE Healthcare).

Helicase and RNA degradation assays

For the decay timecourses in Figure 38B final concentrations were 20 nM exosome, 22 nM or 0 nM Mtr4, 1 mM ATP or AMP-PNP, 0.5 U/µL RNAse inhibitor (New England Biolabs), and 10 nM RNA substrate in RNA decay buffer (20 mM HEPES-KOH pH 7.5, 50 mM potassium acetate, 1.1 mM magnesium acetate, 2.5 mM DTT, 0.01 % IGEPAL). 50 mM ATP and AMP-PNP stock solutions were adjusted to pH 7.0 with KOH. Exosome complexes were incubated on ice with Mtr4 (or buffer) at 2 µM and 2.2 µM, respectively, in RNA decay buffer for 1 hour prior to diluting 1:10 in RNA decay buffer and initiating the reaction. A mix containing RNA, ATP or AMP-PNP, and RNAse inhibitor (all at 1.1x final concentration) was incubated at 20°C for 5 min prior to initiation with 1/10 volume of 200 nM enzyme mix. Reactions were quenched after the indicated incubation times by adding 10 µL of reaction to 5 µL of stop mix (0.3 % w/v SDS, 30 mM EDTA pH 8.0, 3 U/mL proteinase K [New England Biolabs]) followed by proteinase K digestion at 37°C for 1 hour and flash freezing in liquid nitrogen for storage at -80°C. 3.5 µL sample was loaded per lane and run for 45 minutes on 4-20% acrylamide-TBE gels using 0.5x TBE as running buffer. Helicase assays used the exact same protocol except a different substrate was used (see above) and the exosome subunits were inactive in all reactions. For the reactions in Figure 39, reaction conditions and substrate are described in Chapter 2 Materials and Methods.

RNA loading pilot experiments

Catalytically inert Exo13tr complexes were mixed with a 1.1-fold molar excess of Mtr4, diluted to 10 μM in Superdex buffer (20 mM HEPES-KOH pH 7.5, 100 mM KOAc, 0.5 mM $Mg(OAc)_2$, 2.5 mM DTT) and incubated on ice for an hour. During this incubation, four 10x initiation mix containing either 20 mM ATP-KOH pH 7.0, 20 mM AMP-PNP-KOH pH 7.0, ATP plus 4 μM DNA Trap (described above), or ATP plus 2 μM RPA were prepared in reaction buffer plus 20 mM $Mg(OAc)_2$. Then, 11 μL of the Mtr4/exosome mix or buffer (for the substrate only reaction) was mixed with 480 μL of reaction buffer and 2.75 μL of a 40 μM stock of the tripartite substrate with v = 20 nt or buffer (for the –RNA reaction) and incubated at room temperature for 5 minutes. Then, 55 μL of one of the initiation mixes described above was added and the reaction was incubated at room temperature for 40 min. Final concentrations were 0 or 200 nM Exo13tr, 0 or 220 nM Mtr4, 0 or 2 mM ATP-Mg, 0 or 2 mM AMP-PNP-Mg, 0 or 200 nM substrate, 0 or 400 nM DNA trap, and 0 or 200 nM RPA. The reaction was then centrifuged and the supernatant loaded onto a Superdex200 increase 10/300 GL column. 500 μL fractions were collected and the three fractions surrounding the centroid of each peak were analyzed for protein and nucleic acid content. For the nucleic acid gels, samples were treated with proteinase K as above prior to PAGE analysis. SYBR Gold staining was performed per manufacturer's recommendations, and gels were imaged using a Typhoon FLA 9500 instrument.

Protein crosslinking

Prior to crosslinking, proteins were desalted using Bio-Spin 6 columns (Bio-Rad) that had been equilibrated in 20 mM HEPES-KOH pH 7.5, 100 mM KOAc, 1 mM $Mg(OAc)_2$, 0.5 mM ADP and 0.05 mM TCEP-HCl. Complexes were then diluted

to 550 nM and treated with 1/10th volume of a 10x solution of $BM(PEG)_3$ (Thermo Scientific. Prepared fresh in buffer from powder directly before the experiment) or glutaraldehyde and incubated at 4°C for 1 hr for $BM(PEG)_3$ or 25°C for 15 min for glutaraldehyde. The $BM(PEG)_3$ reaction was quenched by the addition of DTT to 2.5 mM and the glutaraldehyde reaction was quenched with the addition of Tris-Cl pH 8.0 to 50 mM. The reactions were analyzed by SDS-PAGE using either SYPRO Ruby staining or Western blotting as previously (see Chapter 2, Materials and Methods).

Sample preparation for cryo-EM

Catalytically inert Exo13tr complexes were mixed with a 1.1-fold molar excess of Mtr4, diluted to 5 µM in loading buffer (20 mM HEPES-KOH pH 7.5, 100 mM KOAc, 0.5 mM $Mg(OAc)_2$, 25 µM TCEP-HCl) and incubated on ice for an hour. During that time, a 10x initiation mix containing 20 mM ATP, 4 µM DNA trap, and 20 mM $Mg(OAc)_2$ in loading buffer was prepared. Then, 40 µL of the Mtr4-exosome mixture was mixed with 1140 µL of loading buffer and 18.6 µL of a 7 µM tripartite substrate with v = 38 nt and incubated at room temperature for 10 min. After this incubation, 130 µL of initiation mix was added and the reaction incubated at room temperature for 50 minutes. Final concentrations were 150 nM Exo14tr, 100 nM substrate, 400 nM DNA trap oligo, and 2 mM ATP-Mg. During this incubation, a Biospin 6 column was equilibrated in loading buffer containing an additional 1 mM $Mg(OAc)_2$ and 1 mM ADP (ADP buffer). Once the reaction was finished, the sample was concentrated to 50 µL in a Microcon YM-30 (Millipore) and buffer exchanged into ADP buffer using the Biospin column. At this point the concentration of the complex was determined by a Bradford assay and the protein diluted to 600 nM with ADP buffer. 16 µL of this solution was aliquoted into a microcentrifuge tube and 2 µL of a 1 mM $BM(PEG)_3$ solution in ADP buffer was added. The sample was incubated

at 4°C for 1 hr to crosslink and was then quenched by the addition of 2 μL of a 25 mM DTT, 0.1% w/v CHAPSO solution in ADP buffer. This resulted in the following final concentrations for the sample: 500 nM complex, 100 μM quenched $BM(PEG)_3$, 0.01% w/v CHAPSO, 2.5 mM DTT, 1.0 mM ADP-Mg in loading buffer.

4 μL of this solution was applied to a UltrAufoil 300 mesh R1.2/1.3 grid (EMS) that had been glow discharged at 15 mA for 1 minute. After a 30 second wait time, the grid was blotted with Whatman paper for 2.5 seconds at 95% humidity and plunged into liquid nitrogen-cooled liquid ethane using a Vitrobot Mark IV (FEI).

Data Collection and processing

Screening and data collection were performed at Memorial Sloan Kettering on the Titan Krios 300 kV microscope equipped with a Gatan K2 Summit direct detector. Data collection was performed in super-resolution mode using a target defocus range of -1.5 to -2.5 μm. 10 second exposures were fractionated to into 50 frames, and the collection resulted in a total of 1483 micrographs with a pixel size of 1.089 Å. After motion correction, pruning via max resolution cutoff from CTF estimations, and manual pruning 1145 micrographs were included for subsequent analysis (see Results section of this chapter for details).

1000 particles from 50 of these movies (chosen at random) were manually selected, extracted (384 pixel box, 350 pixel background circle), and used to generate 2D class averages as references for auto-picking on these micrographs. The result of this was extracted again and subjected to 2D classification, and selected classes from that were used for auto-picking using all 1145 micrographs (360 Å mask, picking threshold 0.25, 150 Å minimum inter-particle distance).

Ab initio 3D models were generated and refined in CryoSPARC and used as starting models for 3D classification in Relion. Additionally, a round of pruning via

selecting from 2D class averages in CryoSPARC enabled the removal of ~20,000 particles in the early phases (Figure 44). After that and an initial round of 3D classification (Figure 43B), focused 3D classification was performed via modification of established protocols (Bai et al., 2015). PDB 4IFD (which contains Exo9 plus RNA, Dis3 in the through-channel conformation, and the Rrp6 EAR) was docked into 3D volumes from refinements in Relion, and its chains corresponding to the RNA and the Rrp6 EAR were removed. The map was then segmented using Segger in Chimera (Pettersen et al., 2004) and the segments were grouped based on chains in the docked model. For the top density focused classification, all of the segments encompassing 4IFD were merged into a single group, and this was used to generate a synthetic map based on those regions. Then, the segments on the top were grouped and used to make another synthetic map. The first synthetic map, which includes essentially every part of the original volume except for the top, was then used to generate a mask that was subtracted from all of the particles (whose orientations are known) in Relion. Using these subtracted particles, a 3D classification was performed without an image alignment and using a mask generate from the second synthetic map (which only encompasses the top densities). Finally, the desired class is selected from these two, the particles re-extracted, and refined again to proceed on to the next step. Focused classification for Csl4 was performed in essentially the same manner, except all of the segments except those encompassing Csl4 were subtracted from the map and the Csl4 region was used as a mask for alignment-free 3D classification.

CONCLUSIONS

The work presented in this dissertation represents a small but significant step forward in our study of RNA exosome function. Specifically, the findings related to conformation control of Dis3 exonuclease activity, channeling of RNA through Exo9 to Rrp6, and the determinants for helicase-assisted degradation of structured substrates all expand on our already rich understanding of the intrinsically regulated nature of RNA decay by the *S. cerevisiae* RNA exosome.

The speculative analogy made nearly two decades ago between RNA degradation by the exosome and protein degradation by the proteasome has turned out to be apt in more ways than the authors could have anticipated. Certainly, both complexes contain active sites that are shielded from their exteriors, but in addition the exosome-associated RNA helicases mirror proteasome-associated unfoldase activities: both consume ATP to unfold structured substrates and thread them through a narrow channel in the complex for their degradation. Stimulation of the nuclear exosome by the TRAMP-dependent addition of 3' poly(A) tails calls to mind the role of K48-linked poly-ubiquitin chains, added to proteins by the vast network of E2 enzymes and E3 ligases, in protein degradation. Furthermore, misfolded or otherwise aberrant proteins and RNAs can both be targeted for poly-ubiquitination or polyadenylation, respectively, over their healthy counterparts. Finally, these stimulatory activities can be antagonized by de-ubiquitinases for proteins and Rrp6 for nuclear transcripts to protect substrates that may have been mistakenly tagged for degradation. In hindsight, it is not terribly surprising that evolution 'decided' to use the same general strategies for restricting spurious protein or RNA degradation by these complexes, both of which are essential for maintaining normal cellular activities.

While *S. cerevisiae* remains a powerful model system with a vast literature, advances in gene editing and mammalian protein expression systems have already begun to rapidly expand the study of exosome function in metazoans. Leveraging these new tools, investigators will have unprecedented control over these systems, and their work will undoubtedly illuminate how the exosome's activities contribute to normal cellular function as well as disease states in humans. We anticipate that structural and biochemical characterization of exosome complexes using metazoan proteins will be essential to our interpretation of these studies, as it has been for *S. cerevisiae*. Hopefully our work, including use of 3'-3' RNAs to trap nuclear complexes and the methods we've developed for observing helicase-dependent RNA decay, can be applied to these systems and be as useful to other investigators as they have been for us in our study of this remarkable complex.

REFERENCES

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Akiyama, B.M., Eiler, D., and Kieft, J.S. (2016). Structured RNAs that evade or confound exonucleases: function follows form. Current Opinion in Structural Biology *36*, 40–47.

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999a). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. The EMBO Journal *18*, 5399–5410.

Allmang, C., Petfalski, E., Podtelejnikov, A., Mann, M., Tollervey, D., and Mitchell, P. (1999b). The yeast exosome and human PM-Scl are related complexes of 3’ --> 5' exonucleases. Genes & Development *13*, 2148–2158.

Aly, H.H., Suzuki, J., Watashi, K., Chayama, K., Hoshino, S.-I., Hijikata, M., Kato, T., and Wakita, T. (2016). RNA Exosome Complex Regulates Stability of the Hepatitis B Virus X-mRNA Transcript in a Non-stop-mediated (NSD) RNA Quality Control Mechanism. J. Biol. Chem. *291*, 15958–15974.

Andersen, P.R., Domanski, M., Kristiansen, M.S., Storvall, H., Ntini, E., Verheggen, C., Schein, A., Bunkenborg, J., Poser, I., Hallais, M., et al. (2013). The human cap-binding complex is functionally connected to the nuclear RNA exosome. Nat Struct Mol Biol *20*, 1367–1376.

Anderson, J.S.J., and Parker, R. (1998). The 3′ to 5′ degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3′ to 5′ exonucleases of the exosome complex. The EMBO Journal *17*, 1497–1506.

Araki, Y., Takahashi, S., Kobayashi, T., Kajiho, H., Hoshino, S., and Katada, T. (2001). Ski7p G protein interacts with the exosome and the Ski complex for 3’ to 5’ mRNA decay in yeast. The EMBO Journal *20*, 4684–4693.

Assenholt, J., Mouaikel, J., Andersen, K.R., Brodersen, D.E., Libri, D., and Jensen, T.H. (2008). Exonucleolysis is required for nuclear mRNA quality control in yeast THO mutants. Rna *14*, 2305–2313.

Astuti, D., Morris, M.R., Cooper, W.N., Staals, R.H.J., Wake, N.C., Fews, G.A., Gill, H., Gentle, D., Shuib, S., Ricketts, C.J., et al. (2012). Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. Nat Genet *44*, 277–284.

Bai, X.-C., Rajendra, E., Yang, G., Shi, Y., and Scheres, S.H.W. (2015). Sampling the conformational space of the catalytic subunit of human γ-secretase. Elife *4*, 1485.

Barandun, J., Chaker-Margot, M., Hunziker, M., Molloy, K.R., Chait, B.T., and Klinge, S. (2017). The complete structure of the small-subunit processome. Nat Struct Mol Biol *24*, 944–953.

Belair, C., Sim, S., and Wolin, S.L. (2017). Noncoding RNA Surveillance: The Ends Justify the Means. Chem. Rev.

Bonneau, F., Basquin, J., Ebert, J., Lorentzen, E., and Conti, E. (2009). The Yeast Exosome Functionsas a Macromolecular Cage to Channel RNA Substrates for Degradation. Cell *139*, 547–559.

Briggs, M.W., Burkard, K.T., and Butler, J.S. (1998). Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation. J. Biol. Chem. *273*, 13255–13263.

Brouwer, R., Allmang, C., Raijmakers, R., van Aarssen, Y., Egberts, W.V., Petfalski, E., van Venrooij, W.J., Tollervey, D., and Pruijn, G.J. (2001). Three novel components of the human exosome. Journal of Biological Chemistry *276*, 6177–6184.

Brown, J.T., Bai, X., and Johnson, A.W. (2000). The yeast antiviral proteins Ski2p, Ski3p, and Ski8p exist as a complex in vivo. Rna *6*, 449–457.

Brunger, A.T. (2007). Version 1.2 of the Crystallography and NMR system. Nature Protocols *2*, 2728–2733.

Burkard, K.T., and Butler, J.S. (2000). A nuclear 3'-5' exonuclease involved in mRNA degradation interacts with Poly(A) polymerase and the hnRNA protein Npl3p. Molecular and Cellular Biology *20*, 604–616.

Butler, J.S., and Mitchell, P. (2011). Rrp6, Rrp47 and cofactors of the nuclear exosome. Adv. Exp. Med. Biol. *702*, 91–104.

Callahan, K.P., and Butler, J.S. (2008). Evidence for core exosome independent function of the nuclear exoribonuclease Rrp6p. Nucleic Acids Res. *36*, 6645–6655.

Castaño, I.B., Heath-Pagliuso, S., Sadoff, B.U., Fitzhugh, D.J., and Christman, M.F. (1996). A novel family of TRF (DNA topoisomerase I-related function) genes required for proper nuclear segregation. Nucleic Acids Res. *24*, 2404–2410.

Chang, H.-M., Triboulet, R., Thornton, J.E., and Gregory, R.I. (2013). A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. Nature *497*, 244–248.

Chapman, M.A., Lawrence, M.S., Keats, J.J., Cibulskis, K., Sougnez, C., Schinzel,

A.C., Harview, C.L., Brunet, J.-P., Ahmann, G.J., Adli, M., et al. (2011). Initial genome sequencing and analysis of multiple myeloma. Nature *471*, 467–472.

Chekanova, J.A., Gregory, B.D., Reverdatto, S.V., Chen, H., Kumar, R., Hooker, T., Yazaki, J., Li, P., Skiba, N., Peng, Q., et al. (2007). Genome-Wide High-Resolution Mapping of Exosome Substrates Reveals Hidden Features in the Arabidopsis Transcriptome. Cell *131*, 1340–1353.

Chen, C.Y., Gherzi, R., Ong, S.E., Chan, E.L., Raijmakers, R., Pruijn, G.J., Stoecklin, G., Moroni, C., Mann, M., and Karin, M. (2001). AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. Cell *107*, 451–464.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Cheng, Z.-F., and Deutscher, M.P. (2002). Purification and characterization of the Escherichia coli exoribonuclease RNase R. Comparison with RNase II. J. Biol. Chem. *277*, 21624–21629.

Chu, L.-Y., Hsieh, T.-J., Golzarroshan, B., Chen, Y.-P., Agrawal, S., and Yuan, H.S. (2017). Structural insights into RNA unwinding and degradation by RNase R. Nucleic Acids Res. *45*, 12015–12024.

Collart, M.A., and Panasenko, O.O. (2012). The Ccr4–Not complex. Gene *492*, 42–53.

Costello, J.L., Stead, J.A., Feigenbutz, M., Jones, R.M., and Mitchell, P. (2011). The C-terminal region of the exosome-associated protein Rrp47 is specifically required for box C/D small nucleolar RNA 3'-maturation. J. Biol. Chem. *286*, 4535–4543.

Dedic, E., Seweryn, P., Jonstrup, A.T., Flygaard, R.K., Fedosova, N.U., Hoffmann, S.V., Boesen, T., and Brodersen, D.E. (2014). Structural analysis of the yeast exosome Rrp6p-Rrp47p complex by small-angle X-ray scattering. Biochem. Biophys. Res. Commun. *450*, 634–640.

Delan-Forino, C., Schneider, C., and Tollervey, D. (2017). Transcriptome-wide analysis of alternative routes for RNA substrates into the exosome complex. PLoS Genet *13*, e1006699.

Didychuk, A.L., Montemayor, E.J., Carrocci, T.J., DeLaitsch, A.T., Lucarelli, S.E., Westler, W.M., Brow, D.A., Hoskins, A.A., and Butcher, S.E. (2017). Usb1 controls U6 snRNP assembly through evolutionarily divergent cyclic phosphodiesterase activities. Nature Communications 1–13.

Domanski, M., Upla, P., Rice, W.J., Molloy, K.R., Ketaren, N.E., Stokes, D.L., Jensen, T.H., Rout, M.P., and LaCava, J. (2016). Purification and analysis of

endogenous human RNA exosome complexes. Rna *22*, 1467–1475.

Drazkowska, K., Tomecki, R., Stodus, K., Kowalska, K., Czarnocki-Cieciura, M., and Dziembowski, A. (2013). The RNA exosome complex central channel controls both exonuclease and endonuclease Dis3 activities in vivo and in vitro. Nucleic Acids Res. *41*, 3845–3858.

Dziembowski, A., Lorentzen, E., Conti, E., and Séraphin, B. (2007). A single subunit, Dis3, is essentially responsible for yeast exosome core activity. Nature Structural & Molecular Biology *14*, 15–22.

Eckard, S.C., Rice, G.I., Fabre, A., Badens, C., Gray, E.E., Hartley, J.L., Crow, Y.J., and Stetson, D.B. (2014). The SKIV2L RNA exosome limits activation of the RIG-I-like receptors. Nat. Immunol. *15*, 839–845.

Egan, E.D., Braun, C.R., Gygi, S.P., and Moazed, D. (2014). Post-transcriptional regulation of meiotic genes by a nuclear RNA silencing complex. Rna *20*, 867–881.

El-Sagheer, A.H., and Brown, T. (2012). Click Nucleic Acid Ligation: Applications in Biology and Nanotechnology. Acc. Chem. Res. *45*, 1258–1267.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Eser, S., Schnieke, A., Schneider, G., and Saur, D. (2014). Oncogenic KRAS signalling in pancreatic cancer. British Journal of Cancer *111*, 817–822.

Fabre, A., and Badens, C. (2014). Human Mendelian diseases related to abnormalities of the RNA exosome or its cofactors. Irdr *3*, 8–11.

Faehnle, C.R., Walleshauser, J., and Joshua-Tor, L. (2014). Mechanism of Dis3l2 substrate recognition in the Lin28–let-7 pathway. Nature *514*, 252–256.

Falk, S., Bonneau, F., Ebert, J., Kögel, A., and Conti, E. (2017). Mpp6 Incorporation in the Nuclear Exosome Contributes to RNA Channeling through the Mtr4 Helicase. CellReports *20*, 2279–2286.

Falk, S., Weir, J.R., Hentschel, J., Reichelt, P., Bonneau, F., and Conti, E. (2014). The Molecular Architecture of the TRAMP Complex Reveals the Organizationand Interplay of Its Two Catalytic Activities. Mol. Cell *55*, 856–867.

Fasken, M.B., Losh, J.S., Leung, S.W., Brutus, S., Avin, B., Vaught, J.C., Potter-Birriel, J., Craig, T., Conn, G.L., Mills-Lujan, K., et al. (2017). Insight into the RNA Exosome Complex Through Modeling Pontocerebellar Hypoplasia Type 1b Disease

Mutations in Yeast. Genetics *205*, 221–237.

Feigenbutz, M., Garland, W., Turner, M., and Mitchell, P. (2013a). The exosome cofactor Rrp47 is critical for the stability and normal expression of its associated exoribonuclease Rrp6 in Saccharomyces cerevisiae. PLoS ONE *8*, e80752.

Feigenbutz, M., Jones, R., Besong, T.M.D., Harding, S.E., and Mitchell, P. (2013b). Assembly of the Yeast Exoribonuclease Rrp6 with Its Associated Cofactor Rrp47 Occurs in the Nucleus and Is Critical for the Controlled Expression of Rrp47. Journal of Biological Chemistry *288*, 15959–15970.

Frazão, C., McVey, C.E., Amblar, M., Barbas, A., Vonrhein, C., Arraiano, C.M., and Carrondo, M.A. (2006). Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. Nature *443*, 110–114.

Fromm, L., Falk, S., Flemming, D., Schuller, J.M., Thoms, M., Conti, E., and Hurt, E. (2017). Reconstitution of the complete pathway of ITS2 processing at the pre-ribosome. Nature Communications *8*, 1787.

Garland, W., Feigenbutz, M., Turner, M., and Mitchell, P. (2013). Rrp47 functions in RNA surveillance and stable RNA processing when divorced from the exoribonuclease and exosome-binding domains of Rrp6. Rna *19*, 1659–1668.

Gasse, L., Flemming, D., and Hurt, E. (2015). Coordinated Ribosomal ITS2 RNA Processing by the Las1 Complex Integrating Endonuclease, Polynucleotide Kinase, and Exonuclease Activities. Mol. Cell *60*, 808–815.

Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. Nature *425*, 737–741.

Greimann, J.C., and Lima, C.D. (2008). Reconstitution of RNA exosomes from human and Saccharomyces cerevisiae cloning, expression, purification, and activity assays. Meth. Enzymol. *448*, 185–210.

Halbach, F., Reichelt, P., Rode, M., and Conti, E. (2013). The Yeast Ski Complex:Crystal Structure and RNA Channeling to the Exosome Complex. Cell *154*, 814–826.

Halbach, F., Rode, M., and Conti, E. (2012). The crystal structure of S. cerevisiae Ski2, a DExH helicase associated with the cytoplasmic functions of the exosome. Rna *18*, 124–134.

Halevy, A., Lerer, I., Cohen, R., Kornreich, L., Shuper, A., Gamliel, M., Zimerman, B.-E., Korabi, I., Meiner, V., Straussberg, R., et al. (2014). Novel EXOSC3 mutation causes complicated hereditary spastic paraplegia. J. Neurol. *261*, 2165–2169.

Hamill, S., Wolin, S.L., and Reinisch, K.M. (2010). Structure and function of the polymerase core of TRAMP, a RNA surveillance complex. Proc. Natl. Acad. Sci. U.S.a. *107*, 15045–15050.

Han, J., and van Hoof, A. (2016). The RNA Exosome Channeling and Direct Access Conformations Have Distinct In Vivo Functions. Cell Reports *16*, 3348–3358.

Hanson, G., and Coller, J. (2018). Codon optimality, bias and usage in translation and mRNA decay. Nat. Rev. Mol. Cell Biol. *19*, 20–30.

Harigaya, Y., Tanaka, H., Yamanaka, S., Tanaka, K., Watanabe, Y., Tsutsumi, C., Chikashige, Y., Hiraoka, Y., Yamashita, A., and Yamamoto, M. (2006). Selective elimination of messenger RNA prevents an incidence of untimely meiosis. Nature *442*, 45–50.

Henras, A.K., Plisson-Chastang, C., O'Donohue, M.-F., Chakraborty, A., and Gleizes, P.-E. (2014). An overview of pre-ribosomal RNA processing in eukaryotes. WIREs RNA *6*, 225–242.

Heo, I., Joo, C., Kim, Y.-K., Ha, M., Yoon, M.-J., Cho, J., Yeom, K.-H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell *138*, 696–708.

Ho, C.K., Rauhut, R., Vijayraghavan, U., and Abelson, J. (1990). Accumulation of pre-tRNA splicing 2'-3' intermediates in a Saccharomyces cerevisiae mutant. The EMBO Journal *9*, 1245–1252.

Hoskins, J.W., Ibrahim, A., Emmanuel, M.A., Manmiller, S.M., Wu, Y., O'Neill, M., Jia, J., Collins, I., Zhang, M., Thomas, J.V., et al. (2016). Functional characterization of a chr13q22.1 pancreatic cancer risk locus reveals long-range interaction and allele-specific effects on DIS3 expression. Hum. Mol. Genet.

Houseley, J., and Tollervey, D. (2009). The Many Pathways of RNA Degradation. Cell *136*, 763–776.

Jackson, R.A., Wu, J.S., and Chen, E.S. (2016). C1D family proteins in coordinating RNA processing, chromosome condensation and DNA damage response. Cell Div *11*, 2.

Jackson, R.N., Klauer, A.A., Hintze, B.J., Robinson, H., van Hoof, A., and Johnson, S.J. (2010). The crystal structure of Mtr4 reveals a novel arch domain required for rRNA processing. The EMBO Journal *29*, 2205–2216.

Januszyk, K., and Lima, C.D. (2014). The eukaryotic RNA exosome. Current Opinion in Structural Biology *24*, 132–140.

Jia, H., Wang, X., Anderson, J.T., and Jankowsky, E. (2012). RNA unwinding by the

Trf4/Air2/Mtr4 polyadenylation (TRAMP) complex. Proc. Natl. Acad. Sci. U.S.a. *109*, 7292–7297.

Jinek, M., Coyle, S.M., and Doudna, J.A. (2011). Coupled 5' nucleotide recognition and processivity in Xrn1-mediated mRNA decay. Mol. Cell *41*, 600–608.

Johnson, A.W., and Kolodner, R.D. (1995). Synthetic lethality of sep1 (xrn1) ski2 and sep1 (xrn1) ski3 mutants of Saccharomyces cerevisiae is independent of killer virus and suggests a general role for these genes in translation control. Molecular and Cellular Biology *15*, 2719–2727.

Jonas, S., Christie, M., Peter, D., Bhandari, D., Loh, B., Huntzinger, E., Weichenrieder, O., and Izaurralde, E. (2014). An asymmetric PAN3 dimer recruits a single PAN2 exonuclease to mediate mRNA deadenylation and decay. Nat Struct Mol Biol *21*, 599–608.

Kadaba, S., Krueger, A., Trice, T., Krecic, A.M., Hinnebusch, A.G., and Anderson, J. (2004). Nuclear surveillance and degradation of hypomodified initiator tRNAMet in S. cerevisiae. Genes & Development *18*, 1227–1240.

Kalisiak, K., Kulinski, T.M., Tomecki, R., Cysewski, D., Pietras, Z., Chlebowski, A., Kowalska, K., and Dziembowski, A. (2016). A short splicing isoform of HBS1L links the cytoplasmic exosome and SKI complexes in humans. Nucleic Acids Res. gkw862.

Kilchert, C., Wittmann, S., and Vasiljeva, L. (2016). The regulation and functions of the nuclear RNA exosome complex. Nat. Rev. Mol. Cell Biol. *17*, 227–239.

Kilchert, C., Wittmann, S., Passoni, M., Shah, S., Granneman, S., and Vasiljeva, L. (2015). Regulation of mRNA Levels by Decay-Promoting Introns that Recruit the Exosome Specificity Factor Mmi1. CellReports *13*, 2504–2515.

Kim, K., Heo, D.-H., Kim, I., Suh, J.-Y., and Kim, M. (2016). Exosome Cofactors Connect Transcription Termination to RNA Processing by Guiding Terminated Transcripts to the Appropriate Exonuclease within the Nuclear Exosome. J. Biol. Chem. *291*, 13229–13242.

Kimanius, D., Forsberg, B.O., Scheres, S.H., and Lindahl, E. (2016). Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. Elife *5*, e18722.

Knapp, G., Ogden, R.C., Peebles, C.L., and Abelson, J. (1979). Splicing of yeast tRNA precursors: structure of the reaction intermediates. Cell *18*, 37–45.

Kowalinski, E., Kögel, A., Ebert, J., Reichelt, P., Stegmann, E., Habermann, B., and Conti, E. (2016). Structure of a Cytoplasmic 11-Subunit RNA Exosome Complex. Mol. Cell *63*, 125–134.

Kowalinski, E., Schuller, A., Green, R., and Conti, E. (2015). Saccharomyces cerevisiae Ski7 Is a GTP-Binding Protein Adopting the Characteristic Conformation of Active Translational GTPases. Structure *23*, 1336–1343.

Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nat Meth *11*, 319–324.

LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005). RNA Degradation by the Exosome Is Promoted by a Nuclear Polyadenylation Complex. Cell *121*, 713–724.

Lebreton, A., Tomecki, R., Dziembowski, A., and Séraphin, B. (2008). Endonucleolytic RNA cleavage by a eukaryotic exosome. Nature *456*, 993–996.

Lee, N.N., Chalamcharla, V.R., Reyes-Turcu, F., Mehta, S., Zofall, M., Balachandran, V., Dhakshnamoorthy, J., Taneja, N., Yamanaka, S., Zhou, M., et al. (2013). Mtr4-like Protein Coordinates Nuclear RNA Processing for Heterochromatin Assembly and for Telomere Maintenance. Cell *155*, 1061–1074.

Lim, J., Ha, M., Chang, H., Kwon, S.C., Simanshu, D.K., Patel, D.J., and Kim, V.N. (2014). Uridylation by TUT4 and TUT7 marks mRNA for degradation. Cell *159*, 1365–1376.

Lim, J., Giri, P.K., Kazadi, D., Laffleur, B., Zhang, W., Grinstein, V., Pefanis, E., Brown, L.M., Ladewig, E., Martin, O., et al. (2017). Nuclear Proximity of Mtr4 to RNA Exosome Restricts DNA Mutational Asymmetry. Cell *169*, 523–537.e15.

Lima, S.A., Chipman, L.B., Nicholson, A.L., Chen, Y.-H., Yee, B.A., Yeo, G.W., Coller, J., and Pasquinelli, A.E. (2017). Short poly(A) tails are a conserved feature of highly expressed genes. Nat Struct Mol Biol *24*, 1057–1063.

Liu, J.-J., Bratkowski, M.A., Liu, X., Niu, C.-Y., Ke, A., and Wang, H.-W. (2014). Visualization of distinct substrate-recruitment pathways in the yeast exosome by EM. Nat Struct Mol Biol *21*, 95–102.

Liu, J.-J., Niu, C.-Y., Wu, Y., Tan, D., Wang, Y., Ye, M.-D., Liu, Y., Zhao, W., Zhou, K., Liu, Q.-S., et al. (2016). CryoEM structure of yeast cytoplasmic exosome complex. Cell Res. *26*, 822–837.

Liu, Q., Greimann, J.C., and Lima, C.D. (2006). Reconstitution, Activities, and Structure of the Eukaryotic RNA Exosome. Cell *127*, 1223–1237.

Lohr, J.G., Stojanov, P., Carter, S.L., Cruz-Gordillo, P., Lawrence, M.S., Auclair, D., Sougnez, C., Knoechel, B., Gould, J., Saksena, G., et al. (2014). Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer Cell *25*, 91–101.

Lorentzen, E., Basquin, J., Tomecki, R., Dziembowski, A., and Conti, E. (2008). Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. Mol. Cell *29*, 717–728.

Losh, J.S., and van Hoof, A. (2015). Gateway Arch to the RNA Exosome. Cell *162*, 940–941.

Losh, J.S., King, A.K., Bakelar, J., Taylor, L., Loomis, J., Rosenzweig, J.A., Johnson, S.J., and van Hoof, A. (2015). Interaction between the RNA-dependent ATPase and poly(A) polymerase subunits of the TRAMP complex is mediated by short peptides and important for snoRNA processing. Nucleic Acids Res. *43*, 1848–1858.

Lowell, J.E., Rudner, D.Z., and Sachs, A.B. (1992). 3'-UTR-dependent deadenylation by the yeast poly(A) nuclease. Genes & Development *6*, 2088–2099.

Lubas, M., Chlebowski, A., Dziembowski, A., and Jensen, T.H. (2012). Biochemistry and Function of RNA Exosomes. Enzymes *31*, 1–30.

Lubas, M., Christensen, M.S., Kristiansen, M.S., Domanski, M., Falkenby, L.G., Lykke-Andersen, S., Andersen, J.S., Dziembowski, A., and Jensen, T.H. (2011). Interaction Profiling Identifies the Human Nuclear Exosome Targeting Complex. Mol. Cell *43*, 624–637.

Lubas, M., Damgaard, C.K., Tomecki, R., Cysewski, D., Jensen, T.H., and Dziembowski, A. (2013). Exonuclease hDIS3L2 specifies an exosome-independent 3’-5’ degradation pathway of human cytoplasmic mRNA. The EMBO Journal *32*, 1855–1868.

Lund, E., and Dahlberg, J.E. (1992). Cyclic 2',3‘-phosphates and nontemplated nucleotides at the 3’ end of spliceosomal U6 small nuclear RNAs. Science *255*, 327–330.

Macias, S., Cordiner, R.A., and Cáceres, J.F. (2013). Cellular functions of the microprocessor. Biochem. Soc. Trans. *41*, 838–843.

Macias, S., Cordiner, R.A., Gautier, P., Plass, M., and Cáceres, J.F. (2015). DGCR8 Acts as an Adaptor for the Exosome Complex to Degrade Double-Stranded Structured RNAs. Mol. Cell *60*, 873–885.

Makino, D.L., Halbach, F., and Conti, E. (2013a). The RNA exosome and proteasome: common principles of degradation control. Nat. Rev. Mol. Cell Biol. *14*, 654–660.

Makino, D.L., Baumgärtner, M., and Conti, E. (2013b). Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. Nature *495*, 70–75.

Makino, D.L., Schuch, B., Stegmann, E., Baumgärtner, M., Basquin, C., and Conti, E. (2015). RNA degradation paths in a 12-subunit nuclear exosome complex. Nature *524*,

54–58.

Malecki, M., Viegas, S.C., Carneiro, T., Golik, P., Dressaire, C.E.M., Ferreira, M.G., and Arraiano, C.I.L.M. (2013). The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. The EMBO Journal *32*, 1842–1854.

Malet, H., Topf, M., Clare, D.K., Ebert, J., Bonneau, F., Basquin, J., Drazkowska, K., Tomecki, R., Dziembowski, A., Conti, E., et al. (2010). RNA channelling by the eukaryotic exosome. EMBO Rep *11*, 936–942.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phasercrystallographic software. J Appl Crystallogr *40*, 658–674.

McIver, S.C., Kang, Y.-A., DeVilbiss, A.W., O'Driscoll, C.A., Ouellette, J.N., Pope, N.J., Camprecios, G., Chang, C.-J., Yang, D., Bouhassira, E.E., et al. (2014). The exosome complex establishes a barricade to erythroid maturation. Blood *124*, 2285–2297.

McIver, S.C., Katsumura, K.R., Davids, E., Liu, P., Kang, Y.-A., Yang, D., and Bresnick, E.H. (2016). Exosome complex orchestrates developmental signaling to balance proliferation and differentiation during erythropoiesis. Elife *5*, 5399.

Meola, N., Domanski, M., Karadoulama, E., Chen, Y., Gentil, C., Pultz, D., Vitting-Seerup, K., Lykke-Andersen, S., Andersen, J.S., Sandelin, A., et al. (2016). Identification of a Nuclear Exosome Decay Pathway for Processed Transcripts. Mol. Cell *64*, 520–533.

Midtgaard, S.F., Assenholt, J., Jonstrup, A.T., Van, L.B., Jensen, T.H., and Brodersen, D.E. (2006). Structure of the nuclear exosome component Rrp6p reveals an interplay between the active site and the HRDC domain. Proc. Natl. Acad. Sci. U.S.a. *103*, 11898–11903.

Miki, T.S., Richter, H., Rüegger, S., and Großhans, H. (2014). PAXT-1 Promotes XRN2 Activity by Stabilizing It through a Conserved Domain. Mol. Cell *53*, 351–360.

Milligan, L., Decourty, L., Saveanu, C., Rappsilber, J., Ceulemans, H., Jacquier, A., and Tollervey, D. (2008). A Yeast Exosome Cofactor, Mpp6, Functions in RNA Surveillance and in the Degradation of Noncoding RNA Transcripts. Molecular and Cellular Biology *28*, 5446–5457.

Mistry, D.S., Chen, Y., and Sen, G.L. (2012). Progenitor Function in Self-Renewing Human Epidermis is Maintained by the Exosome. Cell Stem Cell *11*, 127–135.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997). The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'->5' exoribonucleases. Cell *91*, 457–466.

Mitchell, P. (2010). Rrp47 and the function of the Sas10/C1D domain. Biochem. Soc. Trans. *38*, 1088–1092.

Mitchell, P., Petfalski, E., Houalla, R., Podtelejnikov, A., Mann, M., and Tollervey, D. (2003). Rrp47p is an exosome-associated protein required for the 3' processing of stable RNAs. Molecular and Cellular Biology *23*, 6982–6992.

Molleston, J.M., and Cherry, S. (2017). Attacked from All Sides: RNA Decay in Antiviral Defense. Viruses *9*, 2.

Molleston, J.M., Sabin, L.R., Moy, R.H., Menghani, S.V., Rausch, K., Gordesky-Gold, B., Hopkins, K.C., Zhou, R., Jensen, T.H., Wilusz, J.E., et al. (2016). A conserved virus-induced cytoplasmic TRAMP-like complex recruits the exosome to target viral RNA for degradation. Genes & Development *30*, 1658–1670.

Mukherjee, K., Gardin, J., Futcher, B., and Leatherwood, J. (2016). Relative contributions of the structural and catalytic roles of Rrp6 in exosomal degradation of individual mRNAs. Rna *22*, 1311–1319.

Nagarajan, V.K., Jones, C.I., Newbury, S.F., and Green, P.J. (2013). XRN 5′→3′ exoribonucleases: Structure, mechanisms and functions. Biochim. Biophys. Acta *1829*, 590–603.

Nguyen, D., Grenier St-Sauveur, V., Bergeron, D., Dupuis-Sandoval, F., Scott, M.S., and Bachand, F. (2015a). A Polyadenylation-Dependent 3' End Maturation Pathway Is Required for the Synthesis of the Human Telomerase RNA. CellReports *13*, 2244–2257.

Nguyen, T.A., Jo, M.H., Choi, Y.-G., Park, J., Kwon, S.C., Hohng, S., Kim, V.N., and Woo, J.-S. (2015b). Functional Anatomy of the Human Microprocessor. Cell *161*, 1374–1387.

Ogami, K., Richard, P., Chen, Y., Hoque, M., Li, W., Moresco, J.J., Yates, J.R., Tian, B., and Manley, J.L. (2017). An Mtr4/ZFC3H1 complex facilitates turnover of unstable nuclear RNAs to prevent their cytoplasmic transport and global translational repression. Genes & Development *31*, 1257–1271.

Ohkura, H., Adachi, Y., Kinoshita, N., Niwa, O., Toda, T., and Yanagida, M. (1988). Cold-sensitive and caffeine-supersensitive mutants of the Schizosaccharomyces pombe dis genes implicated in sister chromatid separation during mitosis. The EMBO Journal *7*, 1465–1473.

Otwinowski, Z., and Minor, W. (1997). [20] Processing of X-ray diffraction data collected in oscillation mode. Meth. Enzymol. *276*, 307–326.

Paredes, E., and Das, S.R. (2011). Click chemistry for rapid labeling and ligation of RNA. ChemBioChem *12*, 125–131.

Pefanis, E., Wang, J., Rothschild, G., Lim, J., Chao, J., Rabadan, R., Economides, A.N., and Basu, U. (2014). Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. Nature *514*, 389–393.

Pefanis, E., Wang, J., Rothschild, G., Lim, J., Kazadi, D., Sun, J., Federation, A., Chao, J., Elliott, O., Liu, Z.-P., et al. (2015). RNA Exosome-Regulated Long Non-Coding RNA Transcription Controls Super-Enhancer Activity. Cell *161*, 774–789.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605–1612.

Pirouz, M., Du, P., Munafò, M., and Gregory, R.I. (2016). Dis3l2-Mediated Decay Is a Quality Control Pathway for Noncoding RNAs. CellReports *16*, 1861–1873.

Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M.S., Mapendano, C.K., Schierup, M.H., and Jensen, T.H. (2008). RNA exosome depletion reveals transcription upstream of active human promoters. Science *322*, 1851–1854.

Presnyak, V., Alhusaini, N., Chen, Y.-H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R., et al. (2015). Codon Optimality Is a Major Determinant of mRNA Stability. Cell *160*, 1111–1124.

Proudfoot, N.J. (1989). How RNA polymerase II terminates transcription in higher eukaryotes. Trends in Biochemical Sciences *14*, 105–110.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Meth *14*, 290–296.

Pyle, A.M. (2008). Translocation and unwinding mechanisms of RNA and DNA helicases. Annu Rev Biophys *37*, 317–336.

Reimão Pinto, M.M., Manzenreither, R.A., Burkard, T.R., Sledz, P., Jinek, M., Mechtler, K., and Ameres, S.L. (2016). Molecular basis for cytoplasmic RNA surveillance by uridylation-triggered decay in Drosophila. The EMBO Journal *35*, 2417–2434.

Richter, H., Katic, I., Gut, H., and Großhans, H. (2016). Structural basis and function of XRN2 binding by XTB domains. Nat Struct Mol Biol *23*, 164–171.

Ridley, S.P., Sommer, S.S., and Wickner, R.B. (1984). Superkiller mutations in Saccharomyces cerevisiae suppress exclusion of M2 double-stranded RNA by L-A-HN and confer cold sensitivity in the presence of M and L-A-HN. Molecular and Cellular Biology *4*, 761–770.

Robinson, S.R., Oliver, A.W., Chevassut, T.J., and Newbury, S.F. (2015). The 3’ to 5’

Exoribonuclease DIS3: From Structure and Mechanisms to Biological Functions and Role in Human Disease. Biomolecules *5*, 1515–1539.

Schaeffer, D., Tsanova, B., Barbas, A., Reis, F.P., Dastidar, E.G., Sanchez-Rotunno, M., Arraiano, C.M., and van Hoof, A. (2009). The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. Nat Struct Mol Biol *16*, 56–62.

Schäfer, I.B., Rode, M., Bonneau, F., Schüssler, S., and Conti, E. (2014). The structure of the Pan2-Pan3 core complex reveals cross-talk between deadenylase and pseudokinase. Nat Struct Mol Biol *21*, 591–598.

Schilders, G. (2005). MPP6 is an exosome-associated RNA-binding protein involved in 5.8S rRNA maturation. Nucleic Acids Res. *33*, 6795–6804.

Schilders, G., van Dijk, E., and Pruijn, G.J.M. (2007). C1D and hMtr4p associate with the human exosome subunit PM/Scl-100 and are involved in pre-rRNA processing. Nucleic Acids Res. *35*, 2564–2572.

Schmidt, C., Kowalinski, E., Shanmuganathan, V., Defenouillère, Q., Braunger, K., Heuer, A., Pech, M., Namane, A., Berninghausen, O., Fromont-Racine, M., et al. (2016). The cryo-EM structure of a ribosome-Ski2-Ski3-Ski8 helicase complex. Science *354*, 1431–1433.

Schneider, C., Leung, E., Brown, J., and Tollervey, D. (2008). The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. Nucleic Acids Res. *37*, 1127–1140.

Schneider, C., Kudla, G., Wlotzka, W., Tuck, A., and Tollervey, D. (2012). Transcriptome-wide Analysis of Exosome Targets. Mol. Cell *48*, 422–433.

Schuch, B., Feigenbutz, M., Makino, D.L., Falk, S., Basquin, C., Mitchell, P., and Conti, E. (2014). The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase. The EMBO Journal *33*, 2829–2846.

Schutz, K., Hesselberth, J.R., and Fields, S. (2010). Capture and sequence analysis of RNAs with terminal 2',3'-cyclic phosphates. Rna *16*, 621–631.

Segalla, S., Pivetti, S., Todoerti, K., Chudzik, M.A., Giuliani, E.C., Lazzaro, F., Volta, V., Lazarevic, D., Musco, G., Muzi-Falconi, M., et al. (2015). The ribonuclease DIS3 promotes let-7 miRNA maturation by degrading the pluripotency factor LIN28B mRNA. Nucleic Acids Res. *43*, 5182–5193.

Shchepachev, V., Wischnewski, H., Missiaglia, E., Soneson, C., and Azzalin, C.M. (2012). Mpn1, Mutated in Poikiloderma with Neutropenia Protein 1, Is a Conserved 3′-to-5′ RNA Exonuclease Processing U6 Small Nuclear RNA. Cell Reports *2*, 855–865.

Shchepachev, V., Wischnewski, H., Soneson, C., Arnold, A.W., and Azzalin, C.M. (2015). Human Mpn1 promotes post-transcriptional processing and stability of U6atac. FEBS Lett *589*, 2417–2423.

Shi, Y., Pellarin, R., Fridy, P.C., Fernandez-Martinez, J., Thompson, M.K., Li, Y., Wang, Q.J., Sali, A., Rout, M.P., and Chait, B.T. (2015). A strategy for dissecting the architectures of native macromolecular assemblies. Nat Meth *12*, 1135–1138.

Shukla, S., Schmidt, J.C., Goldfarb, K.C., Cech, T.R., and Parker, R. (2016). Inhibition of telomerase RNA decay rescues telomerase deficiency caused by dyskerin or PARN defects. Nat Struct Mol Biol *23*, 286–292.

Sikorska, N., Zuber, H., Gobert, A., Lange, H., and Gagliardi, D. (2017). RNA degradation by the plant RNA exosome involves both phosphorolytic and hydrolytic activities. Nature Communications *8*, 2162.

Snee, M.J., Wilson, W.C., Zhu, Y., Chen, S.-Y., Wilson, B.A., Kseib, C., O'Neal, J., Mahajan, N., Tomasson, M.H., Arur, S., et al. (2016). Collaborative Control of Cell Cycle Progression by the RNA Exonuclease Dis3 and Ras Is Conserved Across Species. Genetics *203*, 749–762.

Staals, R.H.J., and Pruijn, G.J.M. (2010). The human exosome and disease. Adv. Exp. Med. Biol. *702*, 132–142.

Staals, R.H.J., Bronkhorst, A.W., Schilders, G., Slomovic, S., Schuster, G., Heck, A.J.R., Raijmakers, R., and Pruijn, G.J.M. (2010). Dis3-like 1: a novel exoribonuclease associated with the human exosome. The EMBO Journal *29*, 2358–2367.

Stead, J.A., Costello, J.L., Livingstone, M.J., and Mitchell, P. (2007). The PMC2NT domain of the catalytic exosome subunit Rrp6p provides the interface for binding with its cofactor Rrp47p, a nucleic acid-binding protein. Nucleic Acids Res. *35*, 5556–5567.

Stowell, J.A.W., Webster, M.W., Kögel, A., Wolf, J., Shelley, K.L., and Passmore, L.A. (2016). Reconstitution of Targeted Deadenylation by the Ccr4-Not Complex and the YTH Domain Protein Mmi1. Cell Reports *17*, 1978–1989.

Stuparevic, I., Mosrin-Huaman, C., Hervouet-Coste, N., Remenaric, M., and Rahmouni, A.R. (2013). Cotranscriptional recruitment of RNA exosome cofactors Rrp47p and Mpp6p and two distinct Trf-Air-Mtr4 polyadenylation (TRAMP) complexes assists the exonuclease Rrp6p in the targeting and degradation of an aberrant messenger ribonucleoprotein particle (mRNP) in yeast. J. Biol. Chem. *288*, 31816–31829.

Tafforeau, L., Zorbas, C., Langhendries, J.-L., Mullineux, S.-T., Stamatopoulou, V., Mullier, R., Wacheul, L., and Lafontaine, D.L.J. (2013). The complexity of human

ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. Mol. Cell *51*, 539–551.

Targoff, I.N., and Reichlin, M. (1985). Nucleolar localization of the PM-Scl antigen. Arthritis Rheum. *28*, 226–230.

Taylor, L.L., Jackson, R.N., Rexhepaj, M., King, A.K., Lott, L.K., van Hoof, A., and Johnson, S.J. (2014). The Mtr4 ratchet helix and arch domain both function to promote RNA unwinding. Nucleic Acids Res. *42*, 13861–13872.

Thoms, M., Thomson, E., Baßler, J., Gnädig, M., Griesel, S., and Hurt, E. (2015). The Exosome Is Recruited to RNA Substrates through Specific Adaptor Proteins. Cell *162*, 1029–1038.

Thornton, J.E., Chang, H.-M., Piskounova, E., and Gregory, R.I. (2012). Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). Rna *18*, 1875–1885.

Toh-E, A., Guerry, P., and Wickner, R.B. (1978). Chromosomal superkiller mutants of Saccharomyces cerevisiae. Journal of Bacteriology *136*, 1002–1007.

Tomecki, R., Drazkowska, K., Kucinski, I., Stodus, K., Szczesny, R.J., Gruchota, J., Owczarek, E.P., Kalisiak, K., and Dziembowski, A. (2014). Multiple myeloma-associated hDIS3 mutations cause perturbations in cellular RNA metabolism and suggest hDIS3 PIN domain as a potential drug target. Nucleic Acids Res. *42*, 1270–1290.

Tomecki, R., Kristiansen, M.S., Lykke-Andersen, S.O.R., Chlebowski, A., Larsen, K.M., Szczesny, R.J., Drazkowska, K., Pastula, A., Andersen, J.S., Stepien, P.P., et al. (2010). The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. The EMBO Journal *29*, 2342–2357.

Trotta, C.R., Miao, F., Arn, E.A., Stevens, S.W., Ho, C.K., Rauhut, R., and Abelson, J.N. (1997). The yeast tRNA splicing endonuclease: a tetrameric enzyme with two active site subunits homologous to the archaeal tRNA endonucleases. Cell *89*, 849–858.

Tseng, C.-K., Wang, H.-F., Burns, A.M., Schroeder, M.R., Gaspari, M., and Baumann, P. (2015). Human Telomerase RNA Processing and Quality Control. CellReports *13*, 2232–2243.

Tucker, M., Staples, R.R., Valencia-Sanchez, M.A., Muhlrad, D., and Parker, R. (2002). Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in Saccharomyces cerevisiae. The EMBO Journal *21*, 1427–1436.

Tuma, R.S., Beaudet, M.P., Jin, X., Jones, L.J., Cheung, C.-Y., Yue, S., and Singer, V.L. (1999). Characterization of SYBR Gold Nucleic Acid Gel Stain: A Dye

Optimized for Use with 300-nm Ultraviolet Transilluminators. Analytical Biochemistry *268*, 278–288.

Uchida, N., Hoshino, S.-I., and Katada, T. (2004). Identification of a human cytoplasmic poly(A) nuclease complex stimulated by poly(A)-binding protein. Journal of Biological Chemistry *279*, 1383–1391.

Ustianenko, D., Hrossova, D., Potesil, D., Chalupnikova, K., Hrazdilova, K., Pachernik, J., Cetkovska, K., Uldrijan, S., Zdrahal, Z., and Vanacova, S. (2013). Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. Rna *19*, 1632–1638.

van Dijk, E.L., Chen, C.L., d'Aubenton-Carafa, Y., Gourvennec, S., Kwapisz, M., Roche, V., Bertrand, C., Silvain, M., Legoix-Né, P., Loeillet, S., et al. (2011). XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. Nature *475*, 114–117.

van Hoof, A., and Parker, R. (1999). The exosome: a proteasome for RNA? Cell *99*, 347–350.

van Hoof, A., Staples, R.R., Baker, R.E., and Parker, R. (2000). Function of the Ski4p (Csl4p) and Ski7p Proteins in 3′-to-5′ Degradation of mRNA. Molecular and Cellular Biology *20*, 8230–8243.

Vanacova, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G., and Keller, W. (2005). A New Yeast Poly(A) Polymerase Complex Involved in RNA Quality Control. PLoS Biol *3*, e189–12.

Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S., and Meinhart, A. (2008). The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. Nat Struct Mol Biol *15*, 795–804.

Viegas, S.C., Silva, I.J., Apura, P., Matos, R.G., and Arraiano, C.M. (2015). Surprises in the 3'-end: 'U' can decide too! Febs J. *282*, 3489–3499.

Vincent, H.A., and Deutscher, M.P. (2006). Substrate recognition and catalysis by the exoribonuclease RNase R. J. Biol. Chem. *281*, 29769–29775.

Walker, B.A., Wardell, C.P., Melchor, L., Hulkki, S., Potter, N.E., Johnson, D.C., Fenwick, K., Kozarewa, I., Gonzalez, D., Lord, C.J., et al. (2012). Intraclonal heterogeneity and distinct molecular mechanisms characterize the development of t(4;14) and t(11;14) myeloma. Blood *120*, 1077–1086.

Wan, J., Yourshaw, M., Mamsa, H., Rudnik-Schöneborn, S., Menezes, M.P., Hong, J.E., Leong, D.W., Senderek, J., Salman, M.S., Chitayat, D., et al. (2012). Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration. Nat Genet *44*, 704–708.

Wang, C., Zhu, Y., Bao, H., Jiang, Y., Xu, C., Wu, J., and Shi, Y. (2016). A novel RNA-binding mode of the YTH domain reveals the mechanism for recognition of determinant of selective removal by Mmi1. Nucleic Acids Res. *44*, 969–982.

Wang, H.-W., Wang, J., Ding, F., Callahan, K., Bratkowski, M.A., Butler, J.S., Nogales, E., and Ke, A. (2007). Architecture of the yeast Rrp44 exosome complex suggests routes of RNA recruitment for 3' end processing. Proc. Natl. Acad. Sci. U.S.a. *104*, 16844–16849.

Wang, L., Lewis, M.S., and Johnson, A.W. (2005). Domain interactions within the Ski2/3/8 complex and between the Ski complex and Ski7p. Rna *11*, 1291–1302.

Wasmuth, E.V., and Lima, C.D. (2012a). Exo- and endoribonucleolytic activities of yeast cytoplasmic and nuclear RNA exosomes are dependent on the noncatalytic core and central channel. Mol. Cell *48*, 133–144.

Wasmuth, E.V., and Lima, C.D. (2012b). Structure and Activities of the Eukaryotic RNA Exosome. Enzymes *31*, 53–75.

Wasmuth, E.V., and Lima, C.D. (2017). The Rrp6 C-terminal domain binds RNA and activates the nuclear RNA exosome. Nucleic Acids Res. *45*, 846–860.

Wasmuth, E.V., Januszyk, K., and Lima, C.D. (2014). Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA. Nature *511*, 435–439.

Wasmuth, E.V., Zinder, J.C., Zattas, D., Das, M., and Lima, C.D. (2017). Structure and reconstitution of yeast Mpp6-nuclear exosome complexes reveals that Mpp6 stimulates RNA decay and recruits the Mtr4 helicase. Elife *6*, 213.

Watts, S.G., Crowder, J.J., Coffey, S.Z., and Rubenstein, E.M. (2015). Growth-based determination and biochemical confirmation of genetic requirements for protein degradation in Saccharomyces cerevisiae. J Vis Exp e52428.

Weir, J.R., Bonneau, F., Hentschel, J., and Conti, E. (2010). Structural analysis reveals the characteristic features of Mtr4, a DExH helicase involved in nuclear RNA processing and surveillance. Proc. Natl. Acad. Sci. U.S.a. *107*, 12139–12144.

West, S., Gromak, N., and Proudfoot, N.J. (2004). Human 5’ --> 3’ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature *432*, 522–525.

Wolf, J., and Passmore, L.A. (2014). mRNA deadenylation by Pan2-Pan3. Biochem. Soc. Trans. *42*, 184–187.

Wolf, J., Valkov, E., Allen, M.D., Meineke, B., Gordiyenko, Y., McLaughlin, S.H., Olsen, T.M., Robinson, C.V., Bycroft, M., Stewart, M., et al. (2014). Structural basis for Pan3 binding to Pan2 and its function in mRNA recruitment and deadenylation.

The EMBO Journal *33*, 1514–1526.

Woolford, J.L., and Baserga, S.J. (2013). Ribosome biogenesis in the yeast Saccharomyces cerevisiae. Genetics *195*, 643–681.

Wu, J., and Hopper, A.K. (2014). Healing for destruction: tRNA intron degradation in yeast is a two-step cytoplasmic process catalyzed by tRNA ligase Rlg1 and 5'->3' exonuclease Xrn1. Genes & Development *28*, 1556–1561.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.-C., Dufour, M.-E., Boulay, J., Régnault, B., Devaux, F., Namane, A., Séraphin, B., et al. (2005). Cryptic Pol II Transcripts Are Degraded by a Nuclear Quality Control Pathway Involving a New Poly(A) Polymerase. Cell *121*, 725–737.

Xiang, S., Cooper-Morgan, A., Jiao, X., Kiledjian, M., Manley, J.L., and Tong, L. (2009). Structure and function of the 5'-->3' exoribonuclease Rat1 and its activating partner Rai1. Nature *458*, 784–788.

Yamashita, A., Shichino, Y., Tanaka, H., Hiriart, E., Touat-Todeschini, L., Vavasseur, A., Ding, D.Q., Hiraoka, Y., Verdel, A., and Yamamoto, M. (2012). Hexanucleotide motifs mediate recruitment of the RNA elimination machinery to silent meiotic genes. Open Biology *2*, 120014–120014.

Yang, E., van Nimwegen, E., Zavolan, M., Rajewsky, N., Schroeder, M., Magnasco, M., and Darnell, J.E. (2003). Decay rates of human mRNAs: correlation with functional characteristics and sequence attributes. Genome Res. *13*, 1863–1872.

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. Journal of Structural Biology *193*, 1–12.

Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Meth *14*, 331–332.

Zhou, Y., Zhu, J., Schermann, G.E.Z., Ohle, C., Bendrin, K., Sugioka-Sugiyama, R., Sugiyama, T., and Fischer, T.A.S. (2015). The fission yeast MTREC complex targets CUTs and unspliced pre-mRNAs to the nuclear exosome. Nature Communications *6*, 1–11.

Zinder, J.C., Wasmuth, E.V., and Lima, C.D. (2016). Nuclear RNA Exosome at 3.1 Å Reveals Substrate Specificities, RNA Paths, and Allosteric Inhibition of Rrp44/Dis3. Mol. Cell *64*, 734–745.

Łabno, A., Tomecki, R., and Dziembowski, A. (2016a). Cytoplasmic RNA decay pathways - Enzymes and mechanisms. Biochim. Biophys. Acta *1863*, 3125–3147.

Łabno, A., Warkocki, Z., Kuliński, T., Krawczyk, P.S., Bijata, K., Tomecki, R., and

Dziembowski, A. (2016b). Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs. Nucleic Acids Res. *44*, 10437–10453.

APPENDIX A1: SYNTHESIS OF ASYMMETRIC 3'-3' RNAS

Prior to use of the diazido crosslinking molecule (Figure 9B), our initial strategy for synthesizing 3'-3' RNAs used two differentially 5' modified synthetic RNAs, both of which were purchased from IDT. The first RNA contains a 5' azido modification and is an 11 nt in length, and the second contains a 5' hexynyl group and is of variable length. Copper click chemistry was used to covalently attach these two 5' modifications, resulting in asymmetric RNAs containing two 3' ends (Figure A1.1, panels A and B).

These substrates were tested for complex formation with $Exo12^{Dis3exo\text{-}endo\text{-}/Rrp6exo\text{-}/Rrp47}$ (Figure A1.2) and were used RNase protection assays to determine the ideal length for crystallization (Figure A1.3). The RNase protection assays showed that a catalytically inert $Exo10^{Dis3}$ can protect both 3'-3' and 5'-3' RNAs from degradation by exogenously added catalytically active Rrp6, and that 3'-3' RNAs are trimmed down to the approximately the same size apparently independent of the initial size used. The 3' end of the 5'-3' RNA is presumably engaged by Dis3 and thus is inaccessible towards Rrp6's trimming activity. Furthermore, Rrp6 that lacks the region necessary for interaction with Exo9 (Rrp6ΔNΔC) cannot degrade $Exo10^{Dis3}$ protected 3'-3' RNAs to the same extent as Rrp6 that can associate with Exo9 (Rrp6ΔN), suggesting that these RNAs are indeed being threaded through the central channel to Dis3. Interestingly, these data contrast the channel-independent engagement of 3'-3' RNA by Dis3 observed in the crystal structures (Figures 9D and 27A).

These RNAs were used to determine the preliminary structures of the nuclear exosome in Chapter 1 and the final structure of the Mpp6 exosome presented in Chapter 2 (Figure 27A). Synthesis and purification of these RNAs are described in the Materials and Methods section of Chapter 2.

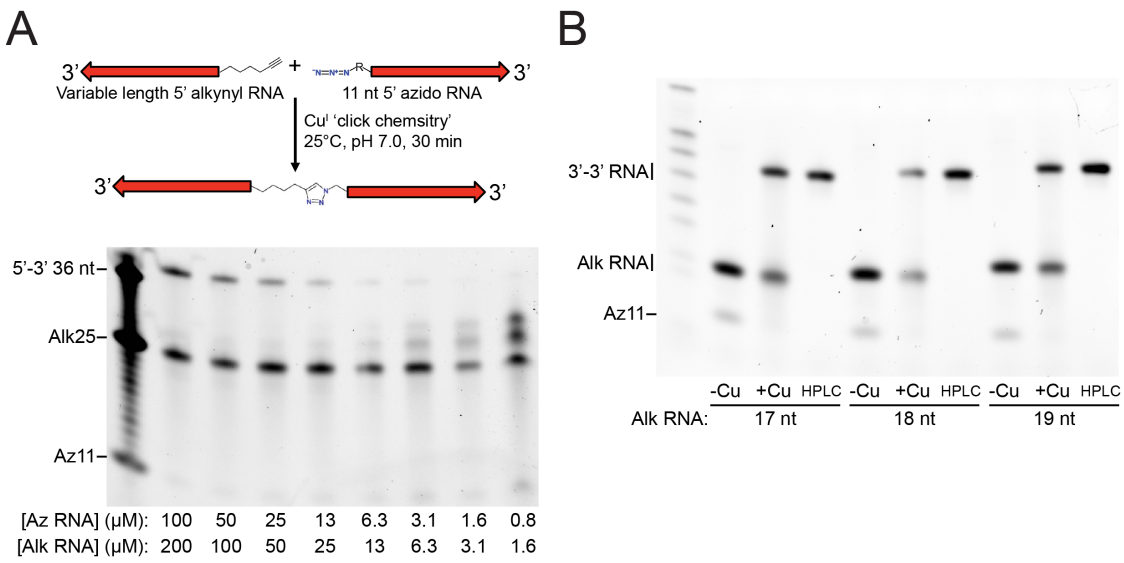


Figure A1.1. Synthesis scheme for asymmetric 3'-3' RNAs. (A) Synthesis scheme for generating 3'-3' RNAs from 5' alkynyl and 5' azido RNAs. RNA is represented as an arrow pointing from 5' to 3'. R represents a proprietary aliphatic linker. Gel is 15% acrylamide TBE-urea and RNA is visualized using SYBR gold stain. Equal ng of total RNA was loaded in each lane. (B) Synthesis of asymmetric 3'-3' RNAs of different lengths. 'HPLC' label indicates reactions that were purified by weak anion exchange as detailed in Chapter 1.

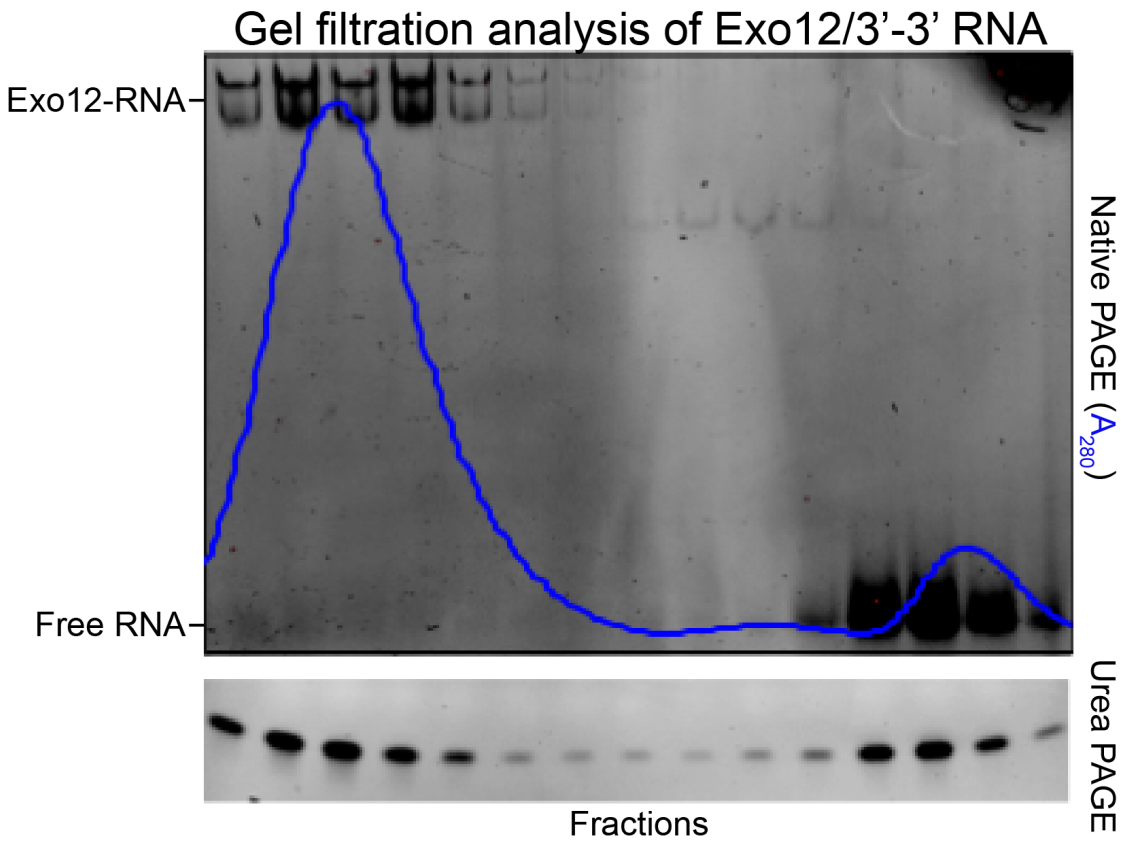


Figure A1.2. Exo12$^{44exo\text{-}endo\text{-}/6exo\text{-}/47}$ forms a complex with an asymmetric 3'-3' 30 nt RNA. A Superdex200 increase 10/300 GL column was run in 20 mM Tris-Cl pH 8.0, 100 mM NaCl, 0.1 mM $MgCl_2$, 0.5 mM TCEP-HCl and 400 µL fractions were collected. Absorbance at 280 nm (A_{280}) is shown as a blue curve roughly aligned to the fractions. Native PAGE is a 4-20% acrylamide TBE gel and Urea PAGE is a 15% acrylamide TBE urea gel. Both are visualized using SYBR Gold staining.

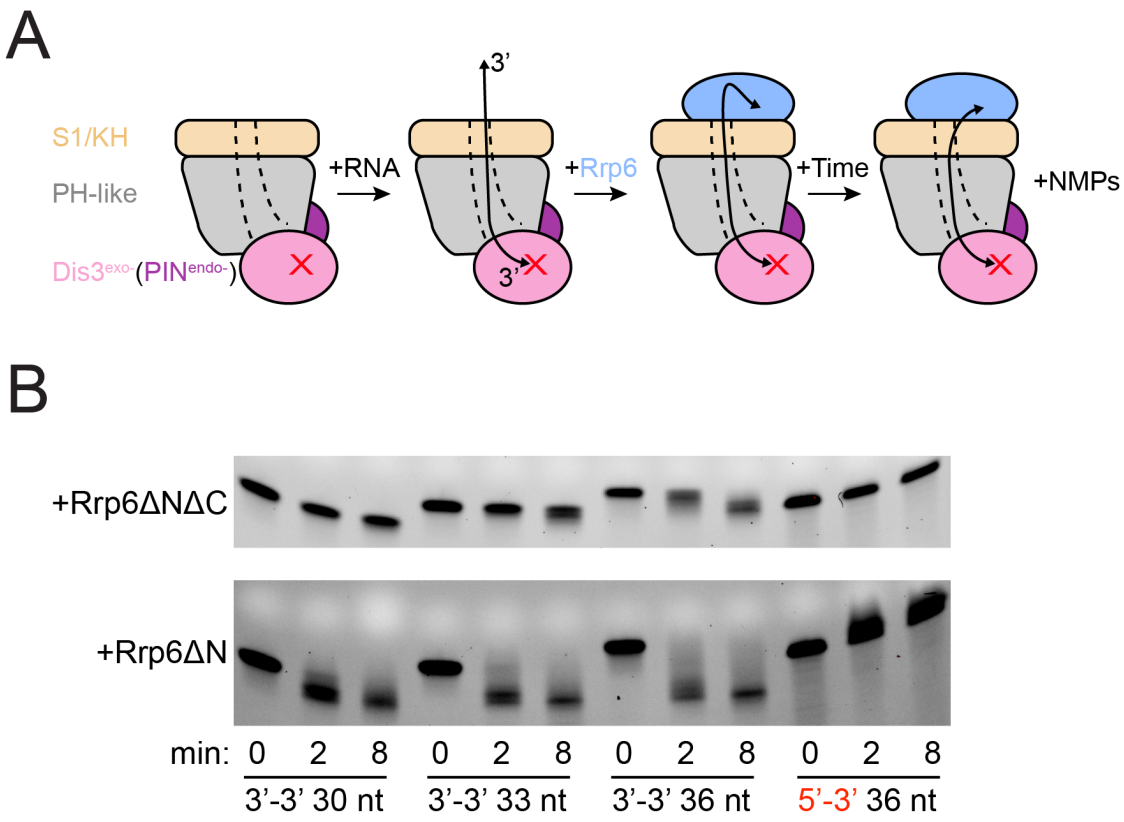


Figure A1.3. Exo10$^{Dis3exo\text{-}endo\text{-}}$ can protect RNA from exogenously added Rrp6. (A) reaction scheme for a Exo10^{Dis3} protection of 3'-3' RNAs from exogenous Rrp6. Side views are shown with exosome components indicated, and the central channel is depicted with dashed lines. RNA is represented as a black line with arrows pointing towards the 3' end. (B) SYBR gold stained 15% acrylamide-urea PAGE of reaction timecourses. Reactions contained 100 nM RNA, 120 nM Exo10$^{Dis3exo\text{-}endo\text{-}}$, 120 nM Rrp6$^{129\text{-}533}$ (Rrp6ΔNΔC) or Rrp6$^{129\text{-}733}$ (Rrp6ΔN) added directly after time = 0 min. RNA and Exo10^{Dis3} were incubated together at 30°C for 5 min at 1.1x final concentration prior to initiation with 1/10th volume of 10x Rrp6. Reactions were conducted in 20 mM Tris-Cl pH 8.0, 50 mM KCl, 0.5 mM $MgCl_2$, 10 mM DTT at 30°C.

APPENDIX A2: 3’ PHOSPHATE RNA DECAY BY DIS3 MUTANTS

In addition to the mutants tested in Figures 24F and 24G, we assayed other mutant Dis3 enzymes for degradation of 3’ phosphate RNA (Figure A2.1). These mutants purified essentially as WT with the exception of the Y595W/R600D double mutant (Figure A2.1 panel B). None of the mutants tested showed differential activity on 3’ phosphate RNA compared to 3’ OH RNA.

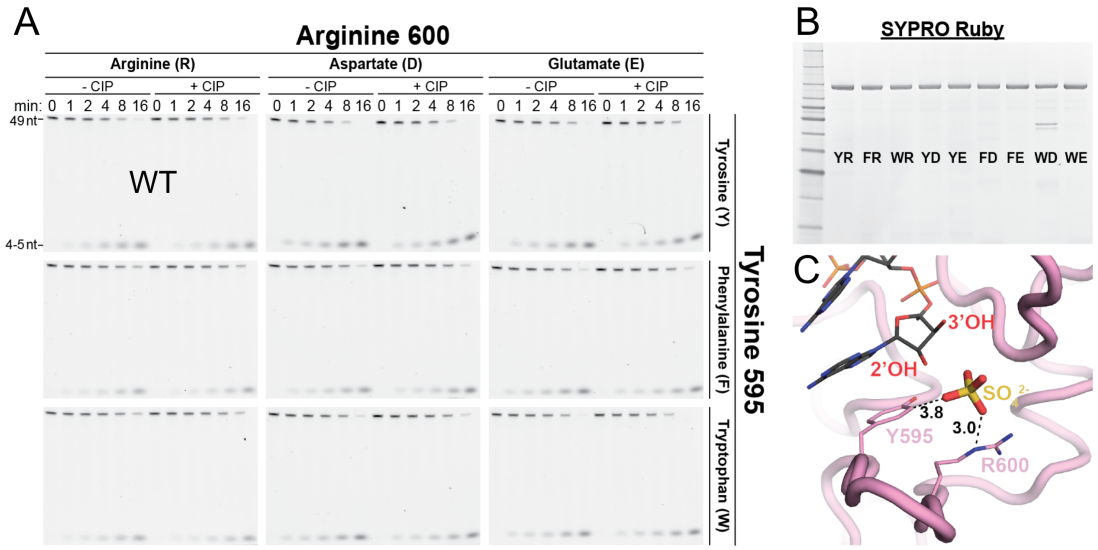


Figure A2.1. Mutants in Arg600 and Tyr595 do not alter Dis3 activity on 3' phosphate RNA relative to 3' OH RNA. (A) RNA decay of 3' phosphate 5' FAM 49 nt AU-rich RNA in the presence or absence of a phosphatase (CIP) by the indicated enzymes. Gels are 15% acrylamide TBE-urea and were imaged for fluorescein signal. (B) SYPRO Ruby stained gels showing the indicated purified mutants. (C) Structure of RNA in Dis3's active site in the nuclear exosome (PDB 5K36) with sidechains of mutated residues shown.

APPENDIX A3: DIS3 PULLDOWN FROM MUTANT YEAST STRAINS

In order to further develop models for Mtr4 recruitment to the exosome in vivo, we transformed a C-terminal TAP-tagged Rrp44 (selected for with a URA marker, from Wasmuth et al., 2012) into selected viable strains from the experiments in Figure 33 and performed IgG pulldowns using lysates from each of the resulting strains in the presence or absence of RNase A (see Chapter 2 Materials and Methods for cell growth and cryo-milling protocols) (Figures A3.1 and A3.2).

One gram of cell powder was resuspended in 3 mL of extraction buffer (50 mM HEPES-NaOH pH 7.5, 100 mM NaCl, 0.5 mM $MgCl_2$, 0.1% v/v IGEPAL ca-630, 5 mM Na-TCEP pH 7.0, 1 mM PMSF, 1x protease inhibitor cocktail [Pierce]), incubated on ice for 10 min, then centrifuged at 20,000 x g for 10 min at 4°C. The supernatant was transferred to fresh tubes, then centrifuged again to further clarify the extract. For the 'input' sample, 40 µL of this supernatant was added to 110 µL of water and 50 µL 4x LDS + BME sample buffer and boiled for 5 min. Then, 1 mL of supernatant was added to 30 µL equilibrated IgG beads (GE Healthcare) and in a separate tube, 1 mL was added to beads + 10 µg RNase A (Qiagen). The tubes were rotated at 4°C for 1 hr, the beads were centrifuged at 500 x g for 30 seconds and the supernatant discarded. The beads were then washed with 900 µL wash buffer (25 mM HEPES-NaOH pH 7.5, 100 mM NaCl, 0.01% v/v IGEPAL ca-630, 1 mM Na-TCEP pH 7.0), centrifuged, then the supernatant discarded. The beads were washed again in a similar fashion, then resuspended in 100 µL TEV solution (50 mM Tris-Cl pH 8.0, 250 mM NaCl, 2 mM Na-TCEP pH 7.0, 0.01 % v/v IGEPAL ca-630, 1 µM TEV protease) and incubated at 4°C overnight. The next morning the beads were agitated, centrifuged, and 90 µl of supernatant was added to 30 µL of 4x LDS + BME sample buffer for the pulldown samples. Lysates and TEV eluates were analyzed for the

presence of exosome components and the TAP tag via Western blotting (Figure A3.1). TEV cleaves between the protein A and CBP portions of the TAG, thus releasing the complex from the beads (Figure A3.1, CBP blot).

This protocol suffered from several technical challenges that made the results essentially uninterpretable. The cryo-milling and/or extraction of from the powder were highly variable, making it very difficult to normalize the pulldown for total protein input (Figure A3.1, Direct Blue stain). Additionally, Mtr4 bound nonspecifically to the beads under the conditions used (Strain N and Lane 1, Figures A3.1 and A3.2 respectively), and at higher salt concentrations (e.g. 150 mM) Mtr4 doesn't bind the exosome in reconstitutions (not shown). Treatment of the lysate with RNase A appeared to mitigate this non-specific binding but not eliminate it (compare lane 1 from Figure 3.2 panels B and C). A positive result from these experiments was that the Rrp47 antibody appears to be functioning well, insofar as only complexes containing the TAP tagged Dis3 and the N-terminus of Rrp6 could pull down Rrp47 (Figure A3.2).

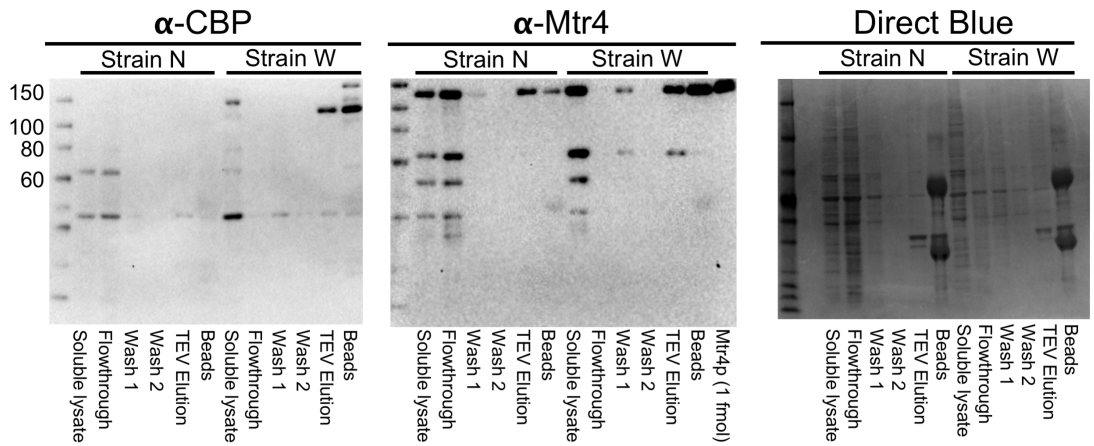


Figure A3.1. Optimization of the pulldown protocol. Strain N and W both contain full length Rrp6 and Mpp6 but strain N lacks the TAP-tagged Dis3. Anti-CBP detects the calmodulin binding protein region of the TAP tag and was purchased from Millipore. Anti-Mtr4 is as described previously (see Materials and Methods, Chapter 2).

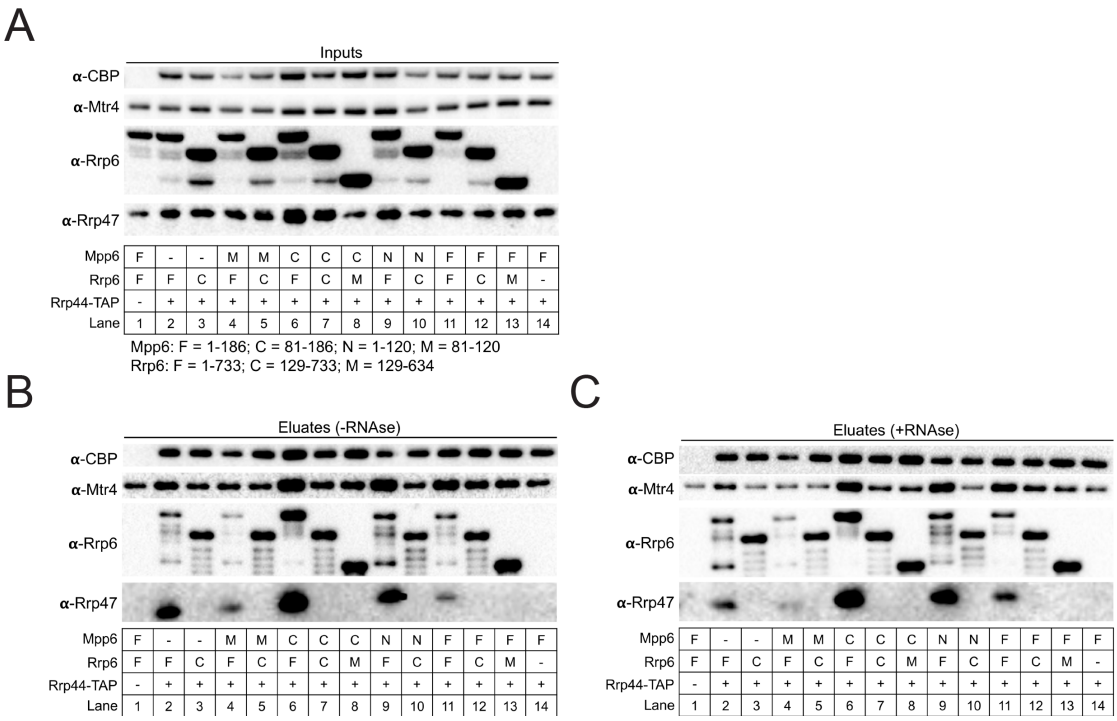


Figure A3.2. Pulldown of TAP-tagged Dis3 in different mutant *S. cerevisiae* strains. Western blotting was performed as previously (Chapter 2, Materials and Methods). Rabbit polyclonal antibodies were raised against calmodulin binding protein (CBP, Millipore), Mtr4 residues 80-1073, Rrp6 residues 129-516, and Rrp47 residues 1-100 and affinity purified. Rrp6 and Mpp6 boundaries are as indicated at the bottom of panel A. (A) Input samples. (B and C) TEV eluates in the absence (B) or presence (C) of RNase A.